



HAL
open science

Modèles paramétriques de processus de branchement uni et multi-types

Amel Ouaari

► **To cite this version:**

Amel Ouaari. Modèles paramétriques de processus de branchement uni et multi-types. Statistiques [math.ST]. Université Montpellier, 2018. Français. NNT : 2018MONTTS109 . tel-02160476

HAL Id: tel-02160476

<https://theses.hal.science/tel-02160476>

Submitted on 19 Jun 2019

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En "Biostatistique"

École doctorale I2S : Information Structures Systèmes

Unité de recherche Biostatistique et Processus Spatiaux (BioSP)
INRA- Avignon (Domaine Saint-Paul)

Modèles paramétriques de processus de branchement uni et multi-types

Présentée par Amel OUAARI
Le 11 Juillet 2018

Sous la direction de Rachid SENOUSSE

Devant le jury composé de

M. Rachid SENOUSSE, Directeur de recherche, INRA

M. Azzedine BENCHETTAH, Professeur, Université Bedji Mokhtar, Annaba, Algérie

Mme Delphine BLANKE, Professeur, Université de Avignon

M. Jean-Noel BACRO, Professeur, Université Montpellier II

Mme Valérie GIRARDIN, Maître de Conférence, Université de Caen

Mme Sophie P ÉNISSON, Maître de Conférence, Université Paris-Est Créteil

Directeur de thèse

Rapporteur

Rapporteur

Examineur

Examinatrice

Examinatrice



UNIVERSITÉ
DE MONTPELLIER

Résumé

L'objet de cette thèse concerne la proposition de modèles paramétriques des processus de branchement uni et multi-types. Nous mettons en valeur l'intérêt de la théorie des processus de branchement et du nécessaire développement de différents outils et concepts propres à leur application dans plusieurs domaines. Pour cela, nous commençons par rappeler quelques définitions et résultats de la théorie des processus de branchement uni et multi-types, et ce en temps discret comme en temps continu. On se consacre par la suite au développement méthodologique de ces modèles.

Dans la deuxième partie de ce mémoire, nous étudions seulement l'évolution d'une seule population en temps continu, et présentons quelques familles de lois paramétriques, associées à des processus de branchement homogènes particuliers. Des méthodes récursives de calcul, ainsi que des propriétés pertinentes, concernant ces distributions de probabilité, sont dérivées des fonctions génératrices satisfaisant certaines équations aux dérivées partielles linéaires précisées. Les familles proposées seront utiles à la modélisation de systèmes plus cohérents en dynamique de populations, puisqu'on y montre que les hypothèses usuelles de distributions de Poisson ne peuvent être argumentées.

Dans la troisième partie, nous étudions le comportement de l'évolution de plusieurs populations en interaction. Nous y présentons aussi des modèles paramétriques de lois, associés à des processus de branchement multi-types en temps continu et homogènes en temps. Nous considérons ensuite un modèle particulier, où une population "mère donneuse" autonome alimente en individus K populations filles, qui sont, elles, en interaction. Ce modèle est bien adapté à l'étude des systèmes dynamiques des populations en interaction qui reste à la fois simple, mais riche en variétés de comportement. L'étude du système multi-types se fait via l'évolution des fonctions génératrices de la loi multidimensionnelle des effectifs. Pour cela, utilisant les équations différentielles ordinaires et aux dérivées partielles, nous établissons les équations implicites des distributions temporelles et multidimensionnelles, et discutons des méthodes analytiques ou numériques de leur résolution. Nous développons ensuite des exemples de modèles et en particulier celui concernant 3 et 4 populations.

En conclusion, nous argumentons la pertinence de cette approche, et l'interprétation des paramètres, qui sont d'un grand intérêt pour le développement de méthodes d'inférence statistique, pour de nombreux domaines d'applications.

Mots clés : Processus de branchement uni-type et multi-types en temps continu, fonction génératrice, systèmes dynamiques, équations différentielles ordinaires, équations aux dérivées partielles.

Abstract

This thesis aims to propose parametric models for single and multi-type branching processes. The importance of the theory of branching processes is pointed out. Hence, developing various tools and specific concepts in several domains is important for applications. For these purposes, we recall some definitions and results of the single-and-multi-type branching processes theory in discrete and continuous case. Afterward, we focus on the methodological development of these models.

In the second part, the evolution of a single population in the continuous case is studied. Then, some parametric distribution families associated to particular branching mechanisms are explored. Recursive computational procedures and relevant properties concerning the associated probability distributions are derived from generating functions that satisfy specified linear partial differential equations. The suggested families are useful for the modeling of systems that are more coherent with population dynamics, contrarily to the usual hypothesis of Poisson distributions, that cannot be argued.

In the third part, the evolution of different populations with interaction is explored. Similarly, some parametric models of homogeneous multi-type branching processes in continuous time are proposed. Afterward, we consider a particular model where an autonomous donor parent population feeds in individuals, K types progeny populations that interact. This model is well adapted to the study of dynamical systems of populations in interaction. This is a simple model, but it has a rich variety of behaviors. The study of such systems is also done regarding the evolution of generating functions of multidimensional individual counting. To achieve such a study, ordinary and partial differential equations are used to establish the implicit equations of temporal and multidimensional distributions. Analytical and numerical methods for equation resolution are then discussed, and examples of particular models are developed.

In conclusion, the relevancy of this approach is argued, considering parameters interpretation in the development of inference methods for the various applied domains.

Keywords : Single-and-multi-type branching processes in continuous time, generating functions, dynamic systems, ordinary differential equations, partial differential equations.

Table des matières

1	Introduction	2
1.1	De l'intérêt des processus de branchement	3
1.2	Un bref historique	4
1.3	Plan de la thèse, contenu des chapitres	10
2	Processus de branchement et construction de modèles paramétriques (uni et multi-types)	12
2.1	Description mathématique des processus de branchement uni-type en temps discret	13
2.2	Processus de Branchement uni-type en temps continu	20
2.3	Quelques exemples simples de branchement	30
2.4	Processus de branchement multi-types homogènes en temps continu . .	35
3	Some prametric families of single-type branching distributions	48
3.1	Introduction	2
3.2	Continuous Time Branching Processes	3
3.3	Fundamental Differential Equations Driving Branching Processes	6
3.4	Examples of Parametric distributions for Branching Processes	13
4	Une famille paramétrique de lois de branchement multi-types	21
4.1	Introduction	22
4.2	Simulation de quelques comportements de branchements multi-types . .	23
4.3	Description du modèle du branchement multi-types	27
4.4	Une famille particulière paramétrique de lois de branchement multi-types	28
4.5	Un exemple de branchement 3-types	41
4.6	Sous-modèle du branchement multi-types mère-filles	48
4.7	Un exemple du sous-modèle : branchement 4-types	52
4.8	Branchement uni-type issu d'un branchement multi-types	60
5	Conclusion et perspectives	64
5.1	Approche statistique par Max-Vraisemblance	65
5.2	Vraisemblance pour des observations ponctuelles	66

5.3 Inversion de la génératrice	67
Bibliographie	69
Références	73

Liste des figures

2.1	Branchement multi-types d'ordre 3.	36
3.1	A. Graph of the primitive $H(z)$ for a cubic branching process. B. A numeric approximation of the function H^{-1} allows the computation of the pgf $G(5, z) = H^{-1}(H(z) + 5)$, $z \in [-1, 1]$, for a cubic branching at a time $t = 5$	18
4.1	Cas explosif de 4 populations indépendantes.	23
4.2	Simulation des tailles dynamiques de 3 populations indépendantes en comportement mixte : explosion extinction.	24
4.3	Simulation des tailles de 3 populations en interactions équitables.	25
4.4	Simulation des tailles de 3 populations en interactions hiérarchique.	26
4.5	Graphe de la solution $x(t)$ de l'EDO homogène du second ordre (4.30) s'amortissant dans le temps vers 0.	39
4.6	Représentation numérique des solutions fondamentales $u_1(t)$ et $u_2(t)$	40
4.7	Représentation numérique de la solution $\mathcal{I}_0(t, \tilde{\mathbf{y}})$ (4.25) de la classe "0".	40
4.8	Simulations de 3 populations en 3 dates.	46
4.9	Histogrammes 2-D des tailles conjointes de populations 2 à 2.	47
4.10	Simulations de 4 populations en 4 dates.	58
4.11	Histogrammes 2-D des tailles conjointes de populations $(X_0(t), X_1(t))$, $(X_0(t), X_3(t))$ et $(X_1(t), X_3(t))$	59
4.12	Histogrammes 2-D des tailles conjointes de 2 populations $(X_0(t), X_2(t))$, $(X_1(t), X_2(t))$ et $(X_2(t), X_3(t))$	60
5.1	Graphe des parties réelles et imaginaires de $G_{X_1}(z)$ (noir), $G_{X_2}(z)$ (jaune), $G_{X_3}(z)$ (vert), $G_{X_4}(z)$ (rouge), $G_{X_5}(z)$ (bleu) pour $z = \frac{1}{2}e^{i\theta}$, $\theta \in [0, 2\pi]$ où Y est le nombre de descendants d'un individu $G_{X_1} = G_Y$	67

Chapitre 1

Introduction

Sommaire

1.1	De l'intérêt des processus de branchement	3
1.2	Un bref historique	4
1.3	Plan de la thèse, contenu des chapitres	10

Dans cette thèse, nous nous intéressons à la construction de modèles paramétriques de processus de branchement uni (PB) et multi-types (PBM). Pour cela, nous commençons par présenter le contexte et les motivations de ce travail. Nous expliquerons la teneur des principaux résultats démontrés dans les chapitres suivants. Dans un but d'inférence statistique, nous avons choisi de mettre en avant les concepts utilisés plutôt que d'insister sur la rigueur et les détails.

1.1 De l'intérêt des processus de branchement

Un processus de branchement est un processus aléatoire, qui décrit l'évolution d'une ou de plusieurs populations ou classes, comme les génotypes et phénotypes en génétique ou le type de particules, électrons, protons etc en physique, chacune composée d'un certain nombre d'individus spécifiques. Chaque individu d'une génération peut donner naissance à de nouveaux individus, selon des règles précisées par des distributions de probabilité spécifiques à la classe. Les processus de branchement concernent le suivi de la taille d'une population, dans laquelle les individus, se reproduisant indépendamment, vivent durant un temps de loi exponentielle de paramètre λ , et à leur fin de vie, sont remplacés par $j \geq 0$ individus avec probabilité p_j .

De ce fait, les processus de branchement sont des approches théoriques bien adaptées pour modéliser des dynamiques que l'on observe dans différents champs appliqués. Ils constituent actuellement une branche importante de la théorie des processus stochastiques. D'une part, les questions mathématiques qu'ils soulèvent sont nombreuses et très riches de nouveaux concepts, et d'autre part ont des domaines d'applications qui ne cessent de s'élargir.

Pour cette raison, les méthodes de processus de branchement sont devenues extrêmement populaires. Elles ont été initiées par Francis Galton, en 1873, qui s'intéressait au problème de transmission des noms de famille. Les processus de branchement interviennent dans plusieurs domaines, tels que :

- la biologie, par exemple dans les réactions en chaîne par polymérase (voir [6, 4, 5, 14]), la sémiotique de la prolifération cellulaire (voir [38, 39]), la spéciation des cellules souches, la théorie de la résistance clonale de cellules cancéreuses,
- l'épidémiologie (voir [25]), pour l'étude de l'évolution et de la propagation de maladies infectieuses,
- la génétique des populations, en considérant l'héritage des allèles, en exemple le processus de branchement multi-types d'ordre 3, correspondant aux génotypes AA, Aa et aa,
- l'astrophysique, où par exemple, les cascades de rayons cosmiques, qui impliquent à la fois des électrons et des photons en interaction, ont été étudiées systématiquement dans [39].

1.2 Un bref historique

Nous présentons ici quelques travaux majeurs établis en processus de branchement. Leur étude remonte à la seconde moitié du XIX^{ème} siècle dans un contexte d'études démographiques.

Les processus de branchement ont débuté en considérant le temps de façon discrète. Ils sont appelés processus de Galton-Watson, ou processus de Bienaymé-Galton-Watson. Ils ont été motivés à l'origine par l'étude des calculs de probabilités d'extinction des noms de famille illustres en Grande-Bretagne. En 1845, Bienaymé [7], probabiliste et statisticien français, a considéré le problème de l'extinction de la famille bien avant que Galton et Watson, mathématiciens et statisticiens britanniques, en 1873 [18, 63] ne le fassent connaître, et en trouvent la solution. Ils ont utilisé pour cela les fonctions génératrices. Grâce à la contribution de Erlang [16] à partir de 1929 puis Guttorp, K Albertsen, JF Steffensen, et E Kristensen en 1995 [21], le problème a été complètement résolu par Fisher [17] en 1999.

Les processus de branchement en temps continu $(X_t, t \in \infty)$, analogues des processus de Galton-Watson en temps discret, sont des processus de Markov forts à valeurs dans \mathbb{N}_* , qui satisfont la propriété de branchement en des temps réels aléatoires. Ils ont été introduits par Jirina en 1958 [28], qui a étendu la propriété de branchement aux processus de Markov, dont l'espace d'état est continu ou discret \mathbb{R}_+ ou \mathbb{N}_* , puis étudiés ensuite par de nombreux auteurs. Watanabe en 1968 et 1969 ([62, 61]) considère des espaces d'état continus plus généraux.

Des travaux concernant certaines propriétés de processus de branchement en temps continu, dûs à Lamperti (1967), caractérisent ces processus comme des limites faibles d'une séquence de processus de Galton-Watson [34]. Dans [35], il obtient le même résultat en opérant un changement de temps aléatoire sur des processus à accroissements indépendants stationnaires. Grimvall en 1974 [20] renforce certains des théorèmes limites de Jirina et Lamperti.

Principaux paramètres des processus de branchement

Commençons par rappeler les principales caractéristiques des processus de branchement dans le cas uni-type. On suppose souvent que la population possède un seul "ancêtre", ce qui se traduit par $X_0 = 1$. La classification courante des processus de branchement fait référence au nombre moyen de descendants $m = \sum_k k p_k$ par individu, p_k étant la probabilité de branchement en k individus.

Le temps d'extinction du processus de branchement est noté $T_0 = \inf\{n : X_n = 0\}$ et $q = P(T_0 < \infty)$ donne la probabilité d'extinction finale. De même, le processus explose au moment $T_\infty = \inf\{n : X_n = \infty\}$ dans le cas uni-type et on a toujours $P(T_\infty \leq \infty) = 1 - q$. Le processus sera dit régulier si $P(T_\infty < \infty) = 1$.

- Il existe deux régimes séparés par une valeur critique du paramètre m :
- Si $m < 1$, le processus est dit “sous-critique”. L’extinction de la population se produit avec probabilité 1 ;
 - Si $m \neq 1$, on a deux cas :
 - Si $m > 1$, le processus est dit “sur-critique”. Alors la probabilité de survie de la population est non nulle ($q < 1$) ;
 - Si $m = 1$, alors le processus est dit “critique”.

Si le processus est sous-critique ou critique, l’extinction est certaine. Si le processus est sur-critique, il y a une probabilité positive que la population survive, et dans ce cas elle explose.

Les processus de branchement non réguliers ont obtenu une attention limitée dans la littérature. Dans les régimes critiques ($m = 1$) et sous-critiques ($m > 1$), l’extinction de la population est presque sûre. Les recherches récentes se sont focalisées sur l’asymptotique de la probabilité de survie ainsi que sur des théorèmes de convergence fonctionnels. De telles publications sont nombreuses, mais on peut se référer aux travaux décrits ci-dessous.

Extinction et explosion des processus de branchement

Nous allons examiner le problème d’extinction pour un processus de branchement uni ou multi-types, traité par de nombreux auteurs.

Les propriétés des processus de branchement par changement de temps aléatoires associés à des processus de Lévy spectralement positifs ont été étudiées par Bingham en 1976 [8]. Il a traité la question des probabilités d’extinction, ainsi que celle de la répartition de la descendance totale d’un PB en temps continu en appliquant la théorie des fluctuations.

Certaines conditions concernant la régularité et l’extinction éventuelle d’un PB ont été obtenues par Chen en 2002 dans [10, 9]. Il a considéré une classe d’équations différentielles associées au PB, qui fournit des outils importants pour l’étude du temps moyen d’extinction et obtient aussi celle de la probabilité d’extinction.

Un PB de Markov pondéré est défini comme un processus de Markov à temps continu à valeurs dans \mathbb{N} , dont la fonction de transition $P(t) = \{p_{i,j}, i, j \in \mathbb{N}_+\}$ satisfait l’équation d’évolution de Kolmogorov : $P'(t) = P(t)Q$, telle que la matrice infinitésimale Q soit de type multiplicatif. Quelques propriétés de base du PB de Markov pondéré ont été traitées par Chen en 2005 [11], pour lesquels, il a donné les critères de régularité et d’unicité de la loi du processus. Il a aussi obtenu le temps moyen d’extinction et le temps moyen d’extinction conditionné par certains événements. Il a également étudié les conditions d’un PB d’explosion de Markov pondéré, et il a évalué le temps moyen.

La probabilité d'extinction d'un processus de branchement multi-types a été étudiée par M.González et al. en 2004 [19]. Leur objectif est de déterminer des conditions suffisantes, pour la probabilité de la survie et de l'extinction du processus, liées à la valeur propre maximale de Perron-Frobenius d'une certaine matrice limite.

Le comportement du temps jusqu'à l'extinction des PB généraux sous-critiques a été traité par Jagers en 2007 [26]. Il a en particulier étudié la taille de la population à mi-chemin de l'extinction. Dans le cas de PB de Markov, il a décrit de façon plus précise le comportement du processus juste avant l'extinction.

Un PB sur-critique général, multi-types, conditionné par l'extinction, se transforme en un processus sous-critique et a été traité par Jagers en 2008 [27], en prouvant que la propriété de branchement vaut aussi pour les processus conditionnés par l'extinction, et en se rappelant que pour un processus de branchement, le processus de reproduction de chaque individu est indépendant de tous les autres individus.

Pakes en 2007 [41] aborde aussi les questions d'extinction/explosion pour une classe de PB Markovien pondérés et détermine les conditions de régularité et la formulation du temps moyen d'extinction dans les cas critique et sur-critique.

La famille des processus de Galton-Watson linéaires-fractionnaires définit la fonction génératrice $f_Y(s) = p_0 + (1+p_0)\frac{p^s}{1-(1-p)s}$ du nombre de descendants $Y \in \mathbb{N}$, représentant la famille des distributions géométriques modifiées, $p_k = (1-p_0)(1-p)^{k-1}p$, où $P(Y = 0) = p_0$, $k \geq 1$ et $p, p_0 \in [0, 1)$. Cette famille de processus de branchement à deux paramètres, ainsi qu'une famille de lois de reproduction à quatre paramètres, ont été étudiées par S.Sagitov en 2015 [49]. Il a montré que cette famille est assez large pour inclure les cas de variance infinie, moyenne infinie, et même des PB non réguliers avec des particules explosives. Il a étudié en détail les propriétés de cette famille, dans certains cas explosifs, où le temps d'explosion est caractérisé asymptotiquement à l'aide d'une distribution de Gumbel.

Le processus de branchement avec immigration (émigration)

Le processus de branchement de Markov avec immigration a été défini d'abord par Sevastyanov en 1957 [53] en termes d'apports populationnistes, qui peuvent se produire pendant de courts intervalles de temps. Il en déduit des théorèmes limites pour le PB de Markov avec immigration, dans le cas où la moyenne des descendants est finie $m = \sum_{j \in \mathbb{N}} jp_j \leq \infty$, p_j étant la probabilité de reproduction pour j descendants et où le défaut d'immigration est nul $d_I = 1 - \sum_{j \in \mathbb{N}} h_j = 0$ avec h_j étant la probabilité d'immigration de taille j . La condition de moment supposée a été levée par Yang en 1972 [65].

L'étude des périodes de présence de vie d'un processus de branchement avec immigration a débuté par Zubkov en 1972 [66]. Il formule les conditions sous lesquelles le temps d'atteinte τ_0 de l'état nul est presque sûrement fini ($P(\tau_0 < \infty) = 1$) pour

un PB critique avec immigration, Vatutin en 1977 [60] obtient des théorèmes de limite pour X_t , étant donné que $\tau_0 > t$.

D'autres théorèmes limites pour la taille de population d'un processus de Bienaymé-Galton-Watson permettant l'immigration ont été obtenus par Pakes en 1979 [43]. Cet article démontre des résultats relatifs aux queues des distributions des probabilités de transition et de la mesure stationnaire.

Un processus de branchement de Markov réglementé par un taux élevé d'émigration, pour éviter une croissance excessive du niveau de population, a été étudié par Chen et Renshaw en 1995 [12]. Ils fournissent aussi un critère d'existence pour de tels modèles, preuve que le processus "stable" correspondant est unique et positif récurrent, en dérivant la fonction génératrice de la distribution limite de la taille de la population.

Dans Li, Chen et Pakes en 2012 [36], la possibilité d'un nombre infini de descendants ou d'immigrants est discutée dans un contexte assez général, pour la construction et l'unicité de la fonction de transition minimale. Cette dernière résout de façon unique le système de Kolmogorov et précise son comportement asymptotique par des théorèmes de limites conditionnelles sur la taille de la population, qui sont donnés dans les cas où la fonction de transition est "instable".

Le processus de branchement avec mutation

Dans le contexte multi-types, un cadre biologiquement pertinent des PB concerne des mutations. Une mutation est un résultat de reproduction lorsque le type de la fille diffère du type de la mère.

Un processus de branchement multi-types non homogène dans le temps, sous l'influence de désastres, distribués selon une loi de Poisson, induisant des taux de mutation variables de particules, a été discuté par Thilaka et al. en 1998 [59]. En particulier, ils ont obtenu l'expression du nombre moyen de particules de chaque type et celle de leur matrice de covariance, sous une forme explicite avec deux types de particules.

Par ailleurs, Kaplan et al. en 1975 [29] ont considéré un processus de branchement uni-type, où les individus en fonction de leur âge sont soumis à des désastres qui se produisent comme un processus de renouvellement. Ils ont étudié le comportement asymptotique de la taille moyenne de la population et fourni des conditions nécessaires et suffisantes pour l'extinction.

Dans un autre contexte, Pakes en 1986 [44] décrit un processus de branchement de Markov, permettant des décrets de population, qui se produisent de manière aléatoire, à un taux proportionnel à la taille de la population. Il en étudie le comportement asymptotique de la probabilité d'extinction et le délai moyen d'extinction et obtient aussi des théorèmes limites pour la taille de la population.

Dans le cas multi-types, on peut avoir des comportements plus complexes du PB. Ainsi, un processus de branchement, régi par une loi de reproduction sous-critique, peut

échapper à l’extinction en raison de mutations, en changeant le taux de reproduction pour le nouveau type de particules avec un régime supercritique. Cela a été démontré par Sagitov en 2009 [50].

Plus récemment, le comportement asymptotique de la population de cellules mutantes a été étudié par Slavtchova et al. en 2017 [54]. Ces auteurs s’intéressent à modéliser l’événement d’apparition d’une cellule mutante d’un type nouveau “succès”, pouvant conduire à la survie de la population. Ils ont développé un algorithme numérique pour résoudre les équations intégrales liées à ce problème, afin d’estimer la répartition du temps d’attente à l’apparence du mutant “succès”. Ensuite, ils décrivent une approche pour la simulation du processus de branchement Bellman-Harris à 2 types. Enfin, ils ont fourni des résultats de simulation sur la répartition du temps d’attente pour atteindre des niveaux élevés du nombre de cellules mutantes.

Processus de branchement en environnement aléatoire

Commençons d’abord par définir ce type de processus de façon plus restreinte.

Un processus de branchement en environnement variable (PBEV) est un cas particulier simple, en l’occurrence si l’environnement est périodique : son étude peut se ramener à l’étude d’un processus de Galton-Watson classique. Ce type de processus de branchement général est lié à la dynamique des individus, qui dépend de l’âge. Un modèle de ces derniers a été formulé par Crump et al. [13] et Ryan [47] en 1968, et pour lequel ils ont démontré des résultats importants dans le cas où le nombre moyen de descendants par individu est suffisamment petit, à savoir $m < 1$.

Les PB multi-types de Markov en temps continu, où les facteurs environnementaux varient de manière périodique, ont aussi été traités par Klein et Macdonald en 1980 [31]. Ils ont étudié, en particulier, le comportement asymptotique de la distribution multivariée des individus de différents types pour le cas sur-critique, décrivant une population croissante. Ils ont prouvé ainsi que le nombre d’individus de différents types converge en probabilité vers un vecteur périodique multiplié par une variable aléatoire (v.a.) dont la distribution dépend de la condition initiale.

Rappelons maintenant la définition d’un processus de branchement dans un environnement aléatoire pour le temps discret. Il s’agit de permettre à chaque génération de choisir une distribution de descendance au hasard, indépendamment des générations précédentes. Si ce choix est périodique, on obtient la classe mentionnée plus haut, et on les définit de la façon suivante. De manière plus générale, si (X_n, E_n) désigne le nombre d’individus à la génération n et le type d’environnement entre les générations n et $n + 1$, et si le nombre de descendants d’un individu à la génération n dans un environnement e de moyenne $m_n(e)$, alors on a :

$$\mathbb{E}(X_{n+1} \mid E_n = e, X_n = k) = k m_k(e).$$

Les résultats principaux des PB en environnement aléatoire, en temps discret ou continu, ont été largement étudiés et développés depuis les années 60. Smith et Wilkinson en 1969 [55] ont étudié, pour le temps discret, la modélisation de la croissance des populations, soumises à un environnement caractérisé, comme une suite iid (E_n) .

La croissance des PB dans un environnement aléatoire, a été étudiée par Tanny en 1977 [57], dans le cadre où on peut avoir à partir de chaque particule du processus une descendance infinie. Ce type de processus soit explose avec un risque exponentiel rapide, soit s'éteint presque sûrement. Il a obtenu de plus un théorème limite pour la probabilité d'extinction en environnement aléatoire stationnaire ergodique.

Très récemment, S. Palau, J.C. Pardo en 2017 [46] ont étudié des probabilités d'explosion et d'extinction des PB en temps continu dans un environnement aléatoire brownien. En particulier, ils ont analysé le comportement à long terme des probabilités de non-explosion et de survie, ainsi que celui du processus conditionné pour une non-extinction et du processus non-conditionné à l'extinction. Enfin, ils ont considéré dans ce même papier le cas d'un processus avec immigration.

Autres questions d'intérêt concernant les processus de branchement

Voici quelques exemples d'importance théorique et pratique qui ont été investis.

EDLER en 1979 [15], s'est intéressé au comportement asymptotique de la distribution de la taille des populations issues de la $K^{\text{ème}}$ génération, en temps t , dans les deux cas : $t \rightarrow \infty$, et $t \rightarrow 0$. Pour le second cas ($t \rightarrow 0$), correspondant à la première génération, il a obtenu des résultats avec des durées de vie non nécessairement exponentiellement distribuées.

Par ailleurs, Küster en 1983 [33] s'est intéressé à des processus de Markov homogènes à valeurs entières, proches des PB uni-types mais où les individus ne sont pas forcément indépendants, ni additifs au niveau générationnel. Il fournit dans ce cadre des résultats asymptotiques presque sûrs concernant les effectifs comme produit d'une fonction déterministe et d'une variable aléatoire unique. Il a généralisé aussi, pour ces processus, des théorèmes de limite centrale et la loi de logarithme itéré obtenue pour des branchements de Markov ordinaires.

Toujours dans le cas de PB homogènes, Pakes (2010) [42] dans le cas critique, conditionne à un temps d'extinction postérieur s'intéresse au comportement du processus lorsque la loi de reproduction est de type α -stable. Il obtient aussi des formules sur la répartition du temps d'extinction asymptotique, par le biais de transformation de Laplace.

Les taux de transition, correspondant à des PB de Markov uni-types en temps continu, sont entièrement déterminés par la fonction génératrice f pour le nombre de descendants d'une seule particule. Dans un souci de généralisation des PBM, S. Sagitov

en 2017 [48] traite le cas des distributions de probabilité "défectueuse" *i.e.* $F(1) < 1$, signifiant qu'avec une probabilité $1 - F(1)$, chaque particule est capable de mettre fin à l'ensemble du processus de branchement : soit le processus s'arrête, soit à un temps T déterminé. Il a considéré ainsi, le cas d'un processus de branchement dit (q, r) -extensible pour lequel il existe q et r tels que si $f(q) = q$ et $f(r) = r$ pour $0 \leq q \leq 1 \leq r < \infty$ avec $q < r$. Avec l'aide de la fonction "génératrice" de la queue $G_t(s) = E(s^{Z_t})$, il a obtenu la répartition du temps de terminaison T conditionnée à une non-extinction $T_0 > T$, et en conclut que les grandes valeurs moyennes du temps de terminaison surviennent dans le cas "presque-critique". Il a obtenu également des résultats asymptotiques analogues pour les processus de branchement sur-critiques.

1.3 Plan de la thèse, contenu des chapitres

L'introduction met en valeur l'intérêt des processus de branchement dans plusieurs domaines. Elle contient un bref historique des débuts jusqu'à plusieurs récents travaux sur ce thème et du développement des différents concepts associés à ces processus.

Nous terminons cette introduction par un résumé des contributions principales ainsi qu'un plan détaillé de notre travail. Ce manuscrit se découpe ainsi en cinq parties d'inégales longueurs.

Le chapitre 2 est consacré au développement méthodologique des modèles des processus de branchement uni ou multi-types et ce en temps discret, comme en temps continu. Nous commençons par présenter les différentes notions, propriétés, définitions et concepts clés de la théorie des processus de branchement uni-types en temps discret. Ensuite, on étudie les équations de branchement uni-type en temps continu qui sont utiles pour la construction de nos modèles, aboutissant à des équations aux dérivées partielles (EDP) et ordinaires linéaires (EDO). On rappelle, outre les notations autour des EDP et des EDO, les définitions et les propriétés des courbes intégrales et des intégrales premières. Ces dernières constituent l'approche classique pour résoudre une EDP linéaire à partir d'une équation différentielle vectorielle, et ses conditions initiales. Cette partie sert au développement du chapitre 3. Ensuite, nous donnons quelques exemples simples d'applications de ces modèles avec plusieurs lois de branchement. Enfin, nous décrivons les outils fondamentaux à la construction de branchement multi-types en temps continu, qui sera à la base du chapitre 4.

Dans cette thèse, nous nous intéressons à divers aspects de processus de branchement et ce dans deux contextes différents, à travers :

- quelques familles paramétriques de distributions en branchement uni-type.
- et une famille paramétrique de lois de branchement multi-types.

Chacun de ces cas correspondra à un chapitre de la thèse. Ces deux chapitres 3 et 4 constituent le cœur de notre travail.

Au cours du chapitre 3, nous présentons quelques familles paramétriques de distributions de probabilité associées à des processus de branchement homogènes particuliers de type unique et en temps continu. Leur simplicité et la pertinence de l'interprétation des paramètres pour les nombreux domaines d'applications sont d'un intérêt certain pour l'inférence statistique. Ces familles sont particulièrement adaptées à l'étude de systèmes dynamiques d'une population. Les calculs et les propriétés pertinentes concernant ces distributions de probabilité sont dérivés des équations des fonctions génératrices aux dérivées partielles linéaires spécifiques.

Dans le chapitre 4, nous présentons des familles paramétriques de lois associées à des processus de branchement multi-types en temps continu et homogènes en temps. Nous considérons un cas particulier (une population mère donneuse autonome et K populations filles en interaction entre elles), utile pour les applications de populations en interaction. On aboutit ainsi à des calculs et à des propriétés concernant leurs lois de probabilité via les fonctions génératrices correspondantes et se rapportant à la résolution de certaines équations aux dérivées partielles linéaires. Une grosse partie de ces équations est résolue analytiquement et l'autre partie par un schéma numérique. Nous donnons ensuite en exemple le cas d'un processus de branchement pour $K = 3$ populations (1 mère-2 filles) du type quadratique, avec quelques simulations illustrant différents types de comportement. Ensuite, nous considérons un cas particulier plus simple du modèle précédent, où chaque population fournit la population qui suit sans en recevoir (exemple de mutations non réversibles). Enfin, nous poursuivons avec l'exemple de ce sous modèle de branchement pour $K = 4$ populations (1 mère-3 filles).

Le dernier chapitre ouvre les perspectives et tire des conclusions de ce travail. Ainsi, nous contribuons au manque flagrant de lois de probabilité spécifiées et paramétrées modélisant les comptages simultanés des individus de plusieurs populations interactives et évoluant dans le temps en dynamique de populations. L'intérêt de développer des méthodes d'inférence statistique propres à ces modèles est de grande actualité. Il existe de nombreuses approches d'estimation paramétriques pour les processus de branchements, mais celles-ci sont pour la plupart basées sur des méthodes de contrastes et de propriétés de martingales [1]. Nous proposons ici de revenir à la méthode du maximum de vraisemblance.

Chapitre 2

Processus de branchement et construction de modèles paramétriques (uni et multi-types)

Sommaire

2.1	Description mathématique des processus de branchement uni-type en temps discret	13
2.1.1	Étude de la fonction génératrice des moments	14
2.1.2	Application au branchement uni-type	15
2.1.3	Probabilité d’extinction	18
2.2	Processus de Branchement uni-type en temps continu	20
2.2.1	Équations de Branchement	20
2.2.2	Générateur d’un processus markovien de sauts homogènes	22
2.2.3	Probabilité d’extinction et moments	27
2.3	Quelques exemples simples de branchement	30
2.3.1	Loi de Bernoulli pour la descendance	30
2.3.2	Loi géométrique pour la descendance	32
2.3.3	Loi de Poisson	34
2.4	Processus de branchement multi-types homogènes en temps continu	35
2.4.1	Outil de base : encore la fonction génératrice	35
2.4.2	Équation de Kolmogorov	36
2.4.3	Probabilité d’extinction et usage de la fonction génératrice des moments	38
2.4.4	EDP, champs de vecteurs et courbes intégrales	43
2.4.5	Application à la résolution de l’EDP “Branchement”	45

Nous présentons dans ce chapitre les processus de branchement uni-type et multi-types. Dans un premier temps, nous commençons par rappeler les définitions et les propriétés de ces processus dans le cas uni et multi-types, et ce en temps discret, comme en temps continu.

2.1 Description mathématique des processus de branchement uni-type en temps discret

En temps discret, les processus de branchements sont appelés processus de Bienaymé-Galton-Watson (BGW). Ils constituent une chaîne de Markov à temps discret, c'est-à-dire une suite de variables aléatoires X_n à mémoire markovienne. X_n représente la taille d'une population d'animaux, de plantes, de cellules ou de gènes au moment de la génération n .

Un processus de branchement modélise l'arbre généalogique aléatoire, issu de la descendance d'un certain nombre d'individus. Un individu naît, vit et donne à un moment donné aléatoire instantanément naissance à des enfants puis meurt. Les outils fondamentaux pour étudier les processus de branchement sont les fonctions génératrices. Elles sont utilisées ainsi pour calculer les probabilités de certains événements associés au processus, ou pour calculer les moments d'intérêt quand ils existent, tels que la moyenne et la variance. Nous présentons les notations et définissons les concepts clés de cette théorie. Elles sont équivalentes aux lois de distributions, mais sont beaucoup plus adaptées à l'étude des phénomènes de branchement.

Définition 1. *Un processus de Bienaymé-Galton-Watson est une suite $X = \{X_n; n > 0\}$ de v.a. à valeurs entières positives décrite récursivement de la manière suivante :*

- À $n = 0$ (population initiale), on a un nombre (aléatoire ou fixe) X_0 d'individus.
- Chaque individu a un temps de vie aléatoire de loi exponentielle, au bout duquel, soit il meurt, soit il se reproduit en Y individus suivant une loi de reproduction $(p_i)_{i \geq 2}$.
- Les individus sont indépendants les uns des autres.

Ainsi si, X_n représente le nombre d'individus de la population au temps $t = n$, en $t = n + 1$, on aura : $X_{n+1} = \sum_{j=1}^{X_n} Y_j$

$$\text{où } Y_j = \begin{cases} 0 & \text{si } j \text{ meurt, avec probabilité } p_0 \\ 1 & \text{si } j \text{ ne meurt pas et se reproduit pas avec probabilité } p_1 \\ k & \text{si } j \text{ meurt donne } k \text{ descendants avec probabilité } p_k, \quad k > 1 \end{cases}$$

Comme les individus sont indifférenciables, on peut aussi convenir que l'individu meurt à chaque pas de temps et donne naissance à $Y_j = k$. Les variables aléatoires Y_j sont

à valeurs entières et sont iid. La somme X_{n+1} , $k \geq 0$ enfants avec probabilités p_k est donc une v.a. discrète à valeurs dans \mathbb{N} : sa loi conditionnellement à $X_n = m$ est la convolution $m^{\text{ième}}$ de la probabilité $P = (p_j)_{j \geq 0}$.

2.1.1 Étude de la fonction génératrice des moments

Dans la théorie des probabilités, la fonction génératrice des moments d'une variable aléatoire à valeurs entières positives ou nulles est une représentation en série de puissances équivalente aux représentations de type Fourier ou Laplace, de la loi de probabilité de cette variable aléatoire.

Définition 2. Si X est une variable aléatoire à valeurs entières positives, la fonction génératrice G_X de X , est définie par la série entière :

$$G_X(z) = \sum_{k=0}^{\infty} \mathbb{P}(X = k)z^k, \quad z \in D(0, 1)$$

qui équivaut à $G_X(z) = \mathbb{E}(z^X)$, où $D(0, 1)$ est le disque complexe $|z| < 1$.

Cette fonction équivaut à la connaissance de la loi de probabilité de X . Comme son nom l'indique, la fonction génératrice donne aisément, les moments de la v.a. quand ils existent et en particulier, l'espérance et la variance.

Théorème 1. Si la variable aléatoire X possède un moment d'ordre k fini, alors on a :

$$\mathbb{E}(X_{(n-1)\dots(n-k+1)}) = G_X^{(k)}(1) = \lim_{\substack{|z| < 1 \\ z \rightarrow 1}} \frac{d^k G_X(z)}{dz^k}$$

De plus, si G_x est deux fois dérivable en $z = 1$, alors on a :

$$\mathbb{E}(X) = G'_X(1) \text{ et } \text{Var}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$$

Pour l'étude des processus de branchement, les fonctions génératrices possèdent la propriété fondamentale suivante de l'addition de v.a. indépendantes.

Proposition 1. Si X et Y sont deux v.a. indépendantes, à valeurs dans \mathbb{N} , alors on a :

$$G_{X+Y}(z) = G_X(z)G_Y(z)$$

Preuve.

il suffit de remarquer que dans ce cas, on a

$$\mathbb{E}(z^{X+Y}) = \mathbb{E}(z^X z^Y) = \mathbb{E}(z^X)E(z^Y) = G_X(z)G_Y(z)$$

Pour mémoire, voici la forme des fonctions génératrices de quelque lois usuelles. Pour tout $z \in D(0, 1) \subset \mathbb{C}$, on a pour :

- une loi de Poisson de paramètre λ , la fonction génératrice $G_X(z) = e^{\lambda(z-1)}$,
- une loi de Bernoulli de paramètre p , la fonction génératrice

$$G_X(z) = 1 + p(z - 1),$$

- une loi Binomiale de paramètre (n, p) et n , la fonction génératrice

$$G_X(z) = (1 - p - pz)^n,$$

- une loi Géométrique à valeurs dans \mathbb{N}^* de paramètre p , la fonction génératrice

$$G_X(z) = \frac{pz}{1 - z(1-p)}.$$

2.1.2 Application au branchement uni-type

Soit $G_Y(z)$ la fonction génératrice de la loi de reproduction Y d'un individu, du processus de branchement X :

$$G_Y(z) = \sum_{k=0}^{\infty} \mathbb{P}(Y = k)z^k.$$

Nous cherchons à calculer, la fonction génératrice $G_n(z)$ du processus X_n , *i.e.* du nombre de descendants à la génération (ou date) n fixée, ainsi que l'espérance et la variance de X_n .

Théorème 2. *La suite de fonctions génératrices $G_{X_n}(z) = G_n(z)$ est définie récursivement de la manière suivante :*

Lorsque $X_0 = 1$ on a : $G_0(z) = z$. Pour $n = 1$, on a $G_1(z) = G(z)$

et

$$G_n(z) = G(G_{X_{n-1}}(z)), \quad (G_{X_n} = G^{o_n}).$$

De plus, si on part de $X_0 = m$ alors la fonction génératrice de X_n s'écrit :

$$G_n^{(m)} = [G_n(z)]^m$$

Preuve.

Nous raisonnons par récurrence.

Le résultat est évident pour $n = 0$ ou 1 . La variable $X_0 = 1$ a pour fonction génératrice $G_{X_0}(z) = z$. De même, par définition de X_1 , la fonction génératrice $G = G_{X_1}$.

Récurrence : Supposons le résultat établi pour tout entier $m \leq n$ avec $X_1 = 1$, c'est à dire $G_m(z) = G_Y^{o_m}$, $\forall m \leq n$

On a :

$$X_{n+1} = \sum_{j=1}^{Y_n} Y_j,$$

avec Y_j : iid de génératrice $G_Y(z)$.

Si $X_0 = 1$ on a

$$\begin{aligned} G_{n+1}(z) &= \mathbb{E}(z^{X_{n+1}}) = \sum_{k=0}^{\infty} (G_Y(z))^k \mathbb{P}(X_n = k) \\ &= G_n(G_Y(z)) = G_{Y_1}^{\circ(n+1)}(z) \end{aligned}$$

et de façon générale si X_0 a pour génératrice $G_0(z)$ alors

$$G_{n+1}(z) = G_0(G_Y^{\circ(n+1)}(z))$$

la condition $X_1 = i$, soit i un entier tel que $\mathbb{P}(X_1 = i) > 0$, alors d'après la proposition 1, la fonction génératrice de X_n est donc $(G_{X_{n-1}})^i$. C'est-à-dire que l'on peut donc écrire pour tout entier i

$$\sum_{k=0}^{\infty} \mathbb{P}(X_n = k \mid X_1 = i) z^k = (G_{X_{n-1}}(z))^i$$

La fonction génératrice de X_n s'écrit donc :

$$\begin{aligned} \sum_{k=0}^{\infty} \mathbb{P}(X_n = k) z^k &= \sum_{k=0}^{\infty} \sum_{\substack{i=0 \\ \mathbb{P}(X_1=i) \neq 0}}^{\infty} \mathbb{P}(X_n = k \mid X_1 = i) \mathbb{P}(X_1 = i) z^k \\ &= \sum_{\substack{i=0 \\ \mathbb{P}(X_1=i) \neq 0}}^{\infty} \left(\sum_{k=0}^{\infty} \mathbb{P}(X_n = k \mid X_1 = i) z^k \right) \mathbb{P}(X_1 = i) \\ &= \sum_{\substack{i=0 \\ \mathbb{P}(X_1=i) \neq 0}}^{\infty} \mathbb{P}(X_1 = i) (G_{X_{n-1}}(z))^i. \end{aligned}$$

Lorsque $\mathbb{P}(X_1 = i) = 0$, on a aussi $\mathbb{P}(X_1 = i) (G_{X_{n-1}}(z))^i = 0$, et donc

$$\sum_{k=0}^{\infty} \mathbb{P}(X_n = k) z^k = \sum_{i=0}^{\infty} \mathbb{P}(X_1 = i) (G_{X_{n-1}}(z))^i.$$

On conclut par récurrence que pour tout n , la variable aléatoire X_n a donc bien pour fonction génératrice $G_{X_n}(z) = G_Y^{\circ n}(z)$.

■

Calculs des deux premiers moments :

Théorème 3. Soit Y tel que, $\mathbb{E}(Y) = m$ et $\text{Var}(Y) = \sigma^2$ alors, si $X_0 = 1$ p.s., on a

$$\text{pour tout } n \geq 2, \mathbb{E}(X_n) = m^n$$

Preuve.

Rappelons que si Y a ses 2 premiers moments finis on a

$$G'(Y) = \mathbb{E}(Y) \quad \text{et} \quad G''(1) = \mathbb{E}(Y^2) - \mathbb{E}(Y)$$

autrement dit $Var(Y) = G''_Y(1) + G'_Y(1) - (G'_Y(1))^2$.

Nous allons procéder par récurrence sur l'entier n , sachant que $\mathbb{E}(x) = G'(1)$ (d'après le théorème 1) et $G_{X_n} = G^{on}$ (d'après le théorème 2), pour tout $n \in \mathbb{N}$.

Pour $n = 0$, $\mathbb{E}(X_0) = \mathbb{E}(1) = 1 = m^0$, $Var(X_0) = 1$
et pour $n = 1$

$$\mathbb{E}(X_1) = \mathbb{E}(Y) = m, \quad Var(X_1) = Var(X) = \sigma^2.$$

Soit $n \in \mathbb{N}$, tel que $\mathbb{E}(X_n) = m^n$. Alors on a :

$$\begin{aligned} \mathbb{E}(X_{n+1}) &= G'_{X_{n+1}}(1) = (G \circ G_{X_n})'(1) \\ &= G'_{X_n}(1)G'(G_{X_n}(1)) = m^{n+1} \quad \text{car} \quad G_n(1) = 1. \end{aligned}$$

■

Théorème 4. *Pour la variance, on opère de même, grâce à :*

$$Var(X_{n+1}) = G''_{n+1}(1) + G'_{n+1}(1) - (G'_{n+1}(1))^2.$$

Pour tout $n \in \mathbb{N}$, on a :

$$Var(X_n) = \begin{cases} n\sigma^2 & \text{si } m = 1 \\ \frac{(\sigma^2 m^{n-1})(1-m^n)}{1-m} & \text{sinon.} \end{cases}$$

Preuve.

En dérivant deux fois G_{X_n} , on obtient :

$$G''_{n+1}(z) = G''_n(G_{n-1}(z))(G'_{n-1}(z))^2 + G'_Y(G_n(z))G''_n(z)$$

et en évaluant en 1, on obtient : $G''_n(1) = G''(1)m^{2(n-1)} + mG''_{n-1}(1)$.

De plus,

$$G''_Y(1) = \sigma^2 - m + m^2 \quad \text{et} \quad G''_{n-1}(1) = \sigma_{n-1}^2 - m^{n-1} - m^{2(n-1)}.$$

Donc

$$Var(X_n) = m^{2(n-1)}\sigma^2 + m\sigma_{n-1}^2.$$

On distingue deux cas :

— Si $m = 1$, $Var(X_n) = \sigma^2 + Var(X_{n-1})$ pour tout $n \in \mathbb{N}$. Ainsi, pour tout $n \in \mathbb{N}$

$$Var(X_n) = n\sigma^2 + Var(X_0) = n\sigma^2.$$

— Si $m \neq 1$, nous raisonnons par récurrence sur la formule annoncée :

Initialisation : $Var(X_0) = 0$, ce qui, dans les deux cas, correspond bien à la formule annoncée.

Hérédité : Soit $n > 2$ tel que

$$Var(X_{n-1}) = \sigma^2 m^{n-2} \frac{1 - m^{n-1}}{1 - m}$$

donc

$$Var(X_n) = \frac{(\sigma^2 m^{n-1})(1 - m^n)}{1 - m}$$

■

2.1.3 Probabilité d'extinction

Un processus de branchement s'éteint s'il existe un $n_0 \in \mathbb{N}$ pour lequel $X_{n_0} = 0$ et on a alors $X_n = 0, \forall n \geq n_0$. Soit q la probabilité que le processus s'éteint. Nous écrivons q_n probabilité que le processus soit éteint exactement à la génération n *i.e.* $q_n = \mathbb{P}(X_n = 0)$. Notez que $q_0 = 0$, en supposant que $X_0 = 1$.

$$\begin{aligned} q_{n+1} &= \mathbb{P}(X_{n+1} = 0) \mathbb{P}(X_{n+1} = 0, X_n = 0) + \mathbb{P}(X_{n+1} = 0, X_n > 0) \\ &= \mathbb{P}(X_n = 0) + \mathbb{P}(X_n > 0) \mathbb{P}(X_{n+1} = 0 \mid X_n > 0) \\ &= q_n + (1 - q_n) \mathbb{P}(X_{n+1} = 0 \mid X_n > 0) \geq q_n. \end{aligned}$$

Donc q_n est une suite croissante, positive, bornée par 1, ce qui implique que :

$$q = \lim_{n \rightarrow \infty} q_n \geq 1 \text{ existe.}$$

Par ailleurs, soit $G_n(z)$ la fonction génératrice des probabilités du processus de branchement.

On a donc $q_n = \mathbb{P}(X_n = 0)G_n(0)$ et q est égale à la plus petite racine non négative de l'équation $s = G(s)$. Comme $G_n(z) = G_Y(G_n(z))$, on a $G_{n+1}(0) = G_Y(G_n(0))$, *i.e.* $q_{n+1} = G_Y(q_n)$ et q vérifie $q = G_Y(q)$. Montrons que, si $m \leq 1$, alors la probabilité d'extinction finale q est 1, *i.e.*

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = 1$$

et si $m > 1$, alors la probabilité d'extinction finale est inférieure à 1, *c.à.d.*

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) < 1.$$

Le processus de Galton-Watson s'appelle sur-critique, critique, sous-critique selon $m > 1, = 1,$ ou < 1 respectivement. Si le processus est sous-critique ou critique, l'extinction est certaine. Mais si le processus est sur-critique, il y a une probabilité positive que la population survive et même explose.

Théorème 5. [32] *On suppose que la variable aléatoire $Y = X_1$ est de carré intégrable avec $E(X_1) = m$ et $Var(X_1) = \sigma^2$. Alors,*

— *si $m = 1$, (Cas critique) on a*

$$\mathbb{P}(X_n > 0) \sim \frac{2}{n\sigma^2}, \quad \text{quand } n \longrightarrow +\infty$$

— *si $m < 1$, (Cas sous-critique) on a*

$$\mathbb{P}(X_n > 0) \sim Cm^n, \quad \text{quand } n \longrightarrow +\infty$$

où C est une constante réelle strictement positive.

2.2 Processus de Branchement uni-type en temps continu

Les processus de branchement en temps continu peuvent être vus comme les analogues en temps et en espace continus des processus de Galton-Watson. Dans le processus de Galton-Watson en temps discret, le temps de vie de chaque particule est l'unité, si on raisonne seulement avec la variable Y du nombre de descendances. La généralisation naturelle en temps continu des processus de branchement markovien est de permettre aux temps de vie d'être des variables aléatoires iid à valeurs réelles positives et par la propriété de Markov d'être forcément distribués de façon exponentielle.

Définition 3. *Un processus de branchement en temps continu $(X_t)_{t \geq 0}$, où X_t représentent un nombre entier d'individus présents au temps t , est défini de façon suivante :*

Au temps $t = 0$, un nombre (aléatoire ou fixe) X_0 d'individus est donné. Ensuite, chaque individu vit un temps exponentiel τ de paramètre λ et à l'instant de sa mort τ , l'individu donne naissance instantanément à i individus avec probabilité p_i . Les temps de vie et les nombres de descendants de chaque individu sont tous indépendants entre eux.

Ainsi, si la population initiale est composée de N individus, les sous-processus issus de ces N ancêtres sont indépendants et de même loi.

2.2.1 Équations de Branchement

Les processus de branchement classiques, étudiés ici sont markoviens, c'est à dire que leur comportement aléatoire dans le futur ne dépend du passé que par l'information donnée par leur état présent. Ce sont aussi des processus de sauts qui peuvent être abordés par la théorie des processus ponctuels marqués.

Définition 4. *Un processus $(X_t; t \geq 0)$ à valeurs dans un espace d'état E est un processus markovien, si pour tout $0 < s < t$, la loi conditionnelle de la variable aléatoire X_t sachant $(X_u; u \leq s)$, ne dépend que de X_s , c'est-à-dire que pour tout $n \in \mathbb{N}$, pour tous $0 \leq t_0 < t_1 < \dots < t_n < s$ et $i_0, i_1, \dots, i_n, i, j \in E$, on a*

$$\mathbb{P}(X_t = j \mid X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n, X_s = i) = \mathbb{P}(X_t = j \mid X_s = i).$$

Introduisons une suite (T_n) croissante de v.a. définie par

$$X_t = X(T_n) \text{ si } T_n \leq t < T_{n+1}.$$

Si E est un espace discret, comme \mathbb{N} ou \mathbb{Q} , tout processus à valeurs dans E est un processus de sauts.

On dit que ce processus markovien est homogène si la probabilité $\mathbb{P}(X_t = j \mid X_s = i)$ ne dépend de s et de t que par la différence $t - s$.

$$\mathbb{P}(X_t = j \mid X_s = i) = P_{i,j}(t - s).$$

Dans le cas particulier où E est au plus dénombrable, la loi est définie par une famille de lois $(P(t))_{t>0}$ données sous forme matricielles vérifiant :

Proposition 2. $P(0) = I_d$ et une relation de semi groupe, appelée **équation de Chapman-Kolmogorov** :

$$P(t + s) = P(t)P(s),$$

où le produit est compris au sens matriciel et est forcément commutatif dans notre cas.

Preuve.

Soit $i, k \in E$. On a

$$\begin{aligned} P_{i,k}(t + s) &= P(X_{t+s} = k \mid X_0 = i) \\ &= \sum_{j \in E} \mathbb{P}(X_t = j, X_{t+s} = k \mid X_0 = i) \\ &= \sum_{j \in E} \mathbb{P}(X_t = j, X_{t+s} = k \mid X_t = j) \mathbb{P}(X_t = j \mid X_0 = i). \end{aligned}$$

par la propriété de Markov, d'où $P(t + s) = P(t)P(s)$, dite de semi-groupe. ■

Dans le cas des processus de branchement, il est plus facile, pour décrire la loi du processus au temps t , d'étudier la famille des fonctions génératrices.

Définition 5. La fonction génératrice d'un processus markovien à valeurs dans \mathbb{N} avec $X(0) = 1$, pour $z \in D(0, 1)$, est défini par :

$$G^{[1]}(t, z) = \mathbb{E} \left(z^{X_t} \mid X_0 = 1 \right) = \sum_{j \geq 0} \mathbb{P}(X_t = j \mid X_0 = 1) z^j = \sum_{j \geq 0} P_{1,j}(t) z^j. \quad (2.1)$$

La probabilités de transition $P(t)$ associée à la fonction génératrice de X_t , satisfait la propriété de branchement, à savoir que pour tout temps t et $z \in D(0, 1)$, pour tout $i \geq 0$, tel que $X(0) = i$,

$$G^{[i]}(t, z) = \sum_{j \geq 0} P_{i,j}(t) z^j = \left(G^{[1]}(t, z) \right)^i = \left(\sum_{j \geq 0} P_{1,j}(t) z^j \right)^i. \quad (2.2)$$

2.2.2 Générateur d'un processus markovien de sauts homogènes

La matrice de transition $P(t)$ qui décrit la loi de $X(t)$ au temps $t \geq 0$ peut être déterminée par sa dérivée à droite de $P(t)$ au temps $t = 0$, appelée générateur infinitésimal du processus. Il permet de décrire la structure du processus en s'aidant de la propriété de Markov forte.

Définition 6. Soit $(P(t))_{t>0}$, le semi-groupe de matrices de transition d'un processus markovien homogène de sauts $\{X_t, t > 0\}$ à valeurs dans E dénombrable. Il existe alors, une matrice $Q = (Q_{i,j}, i, j \in E)$, appelée générateur infinitésimal, qui vérifie

$$Q_{i,j} \geq 0 \quad \text{pour } i \neq j \quad \text{et} \quad Q_{i,i} = - \sum_{\substack{j \in E \\ j \neq i}}^{\infty} Q_{i,j} \leq 0$$

et pour $h \rightarrow 0$

$$P_{i,j}(h) = hQ_{i,j} + o(h) \quad \text{si } i \neq j \tag{2.3}$$

$$P_{i,i}(h) = 1 + hQ_{i,i} + o(h) \quad \text{si } i = j \tag{2.4}$$

Les relations (2.3) et (2.4) signifient que le générateur infinitésimal Q correspond à un opérateur de dérivation de la loi du processus, défini pour toute fonction bornée sur E par :

$$Qf(i) = \lim_{t \rightarrow 0, t > 0} \frac{P_t f(i) - f(i)}{t}$$

Expliquons le lien avec la matrice Q que nous venons d'introduire en utilisant (2.4).

Notation : Pour une fonction bornée $f : E \rightarrow \mathbb{R}$ notée $f \in \mathcal{B}(E, \mathbb{R})$, on définit $P_t f(i) = \mathbb{E}(f(X_t) | X(0) = i)$. Alors on a

$$\begin{aligned} P_t f(i) &= \sum_j P_{i,j} f(j) = P_{i,i} f(i) + \sum_{j \neq i} P_{i,j} f(j) \\ &= (1 + hQ_{i,i} + o(h))f(i) + \sum_{j \neq i} (hQ_{i,j} + o(h))f(j) \\ &= f(i) + h \left(\sum_{j \neq i} Q_{i,j} (f(j) - f(i)) \right) + o(h), \end{aligned}$$

c'est à dire

$$Qf(i) = \sum_{j \neq i} Q_{i,j} (f(j) - f(i)),$$

comme conséquence de la propriété de Markov.

Définition 7. $Q_{i,i}$ s'appelle le taux de sauts du processus dans son état i , et nous avons

$$\mathbb{P}(T_1 > t \mid X_0 = i) = e^{Q_{i,i}t},$$

où T_1 est le premier instant de saut de loi exponentielle de paramètre $-Q_{i,i}$.

Théorème 6. La probabilité de transition d'un processus de branchement $(P_{i,j}(t), t \geq 0)$ est l'unique solution de l'équation de Kolmogorov backward (rétrograde)

$$\partial_t P_{i,j}(t) = \sum_{k=0}^{\infty} Q_{i,k} P_{k,j}(t).$$

$(P_{i,j}(t), t \geq 0)$ est l'unique solution de l'équation de Kolmogorov forward (progressive)

$$\partial_t P_{i,j}(t) = \sum_{k=0}^{\infty} P_{i,k}(t) Q_{k,j}.$$

Preuve. (rapide)

$$\begin{aligned} \partial_t P_{i,j}(t) &= \lim_{h \rightarrow 0} \frac{P_{i,j}(t+h) - P_{i,j}(t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\sum_{k=0}^{\infty} P_{i,k}(h) P_{k,j}(t) - P_{i,j}(t) \right] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\sum_{\substack{k=0 \\ k \neq i}}^{\infty} P_{i,k}(h) P_{k,j}(t) + P_{i,i}(h) P_{i,j}(t) - P_{i,j}(t) \right] \\ &= \sum_{\substack{k=0 \\ k \neq i}}^{\infty} \lim_{h \rightarrow 0} \frac{P_{i,k}(h)}{h} P_{k,j}(t) + \lim_{h \rightarrow 0} \left(\frac{P_{i,i}(h) - 1}{h} \right) P_{i,j}(t) \\ &= \sum_{\substack{k=0 \\ k \neq i}}^{\infty} Q_{i,k} P_{k,j}(t) + Q_{i,i} P_{i,j}(t) = \sum_{k=0}^{\infty} Q_{i,k} P_{k,j}(t). \end{aligned}$$

■

Dans le cas uni-type, la propriété de Markov va se traduire entre autre par le résultat suivant :

Théorème 7. Supposons $X_0 = 1$. Nous avons alors, pour tous $0 < s < t$:

1.

$$G^{[1]}(t+s, z) = G^{[1]}(s, G(t, z)) \tag{2.5}$$

2.

$$\begin{cases} \partial_t G^{[1]}(0, z) = \sum_{k \geq 0} Q_{1,k} [G^{[1]}(t, k)]^k \\ G^{[1]}(0, z) = z \end{cases} \tag{2.6}$$

Preuve.

1. Pour tous $0 < s < t$, on a :

$$\begin{aligned}
 G^{[1]}(t+s, z) &= \mathbb{E}\left(z^{X(s+t)}\right) = \sum_{j \geq 0} \mathbb{P}(X(s) = j) \mathbb{E}(z^{X(s+t)} \mid X(s) = j) \\
 &= \sum_{j \geq 0} \mathbb{P}(X(s) = j) \mathbb{E}(z^{X(s)} \mid X(0) = j) \\
 &= \sum_{j \geq 0} \mathbb{P}(X(s) = j) \left(\sum_{k \geq 0} P_{1,k}(t) z^k \right)^j \\
 &= \sum_{j \geq 0} \mathbb{P}(X(s) = j) (G(t, z))^j.
 \end{aligned}$$

D'où $G^{[1]}(t+s, z) = G^{[1]}(s, G^{[1]}(t, z))$.

2. On a également $G^{[1]}(0, z) = \mathbb{E}(z^{X_0}) = \mathbb{E}(z) = z$.

$$\begin{aligned}
 \partial_t G^{[1]}(t, z) &= \sum_{j \geq 0} \left(\frac{\partial P_{1,j}(t)}{\partial t} \right) z^j \\
 &= \sum_{j \geq 0} \sum_{k \geq 0} Q_{1,k} P_{k,j}(t) z^j \\
 &= \sum_{k \geq 0} Q_{1,k} (G^{[1]}(t, z))^k.
 \end{aligned}$$

Lorsque $t = 0$, on obtient $G^{[1]}(0, z) = z$, d'où

$$\partial_t G^{[1]}(0, z) = \sum_{k \geq 0} Q_{1,k} z^k.$$

■

Formules du générateur pour un PB uni-type :

Proposition 3. *Le générateur du processus de branchement en temps continu est donné par les formules suivantes.*

Si $\lambda = -Q_{1,1} > 0$, alors,

$$Q_{i,i} = -i\lambda, \quad i \geq 0,$$

et pour $j \neq i$:

$$Q_{i,j} = \begin{cases} i\lambda\mu & \text{si } j = i - 1 \\ i\lambda(1 - \mu)p_k & \text{si } j = i + k \end{cases}$$

où $(p_k)_{k \geq 1}$ est la loi du nombre de descendants Y et μ est la probabilité pour un individu de mourir.

Preuve.

Nous savons que

$$Q_{i,i} = \lim_{t \rightarrow 0} \frac{1 - P_{i,i}(t)}{t}.$$

Comme $1 - P_{i,i}(t) = P_i(X(t) \neq i)$, montrons que cette probabilité de saut est proche de $P_i(T_1 < t) = 1 - e^{-i\lambda t}$ au voisinage de $t \sim 0$, où T_1 est l'instant de premier saut.

En effet, si on pose $A = P_i(T_1 < t)$ et $B = P_i(X(t) \neq i)$, si $B \subset A$ alors $P(A - B)$, *i.e.* il y a au moins 2 sauts avant t , un pour quitter i et un pour revenir, d'où

$$P_i(T_1 < t) - P_i(X(t) \neq i) \leq P_i(T_2 < t),$$

or $P_i(T_2 < t) = P_i(T_1 + T_2 - T_1) \leq P_i(T_1 < t)P_i(T_2 - T_1)$ car par construction, T_1 et $T_2 - T_1$ sont indépendants. Nous déduisons que $P_i(T_2 - T_1)$ est d'ordre t^2 , négligeable devant $P_i(T_1 < t)$ pour tout t petit.

$$Q_{i,i} = \lim_{t \rightarrow 0} \frac{1 - P_{i,i}}{t} = \lim_{t \rightarrow 0} \frac{1 - e^{-i\lambda t}}{t} = i\lambda$$

Avec un argument similaire, nous avons également que pour $i \neq j$,

$$\begin{aligned} P_{i,j}(t) &= \mathbb{P}(X_t = j | X_0 = i) \\ &= \mathbb{P}(T_1 < t, X_{T_1} = j | X_0 = i) + \varrho(t) \\ &= \mathbb{P}(T_1 < t | X_0 = i) \mathbb{P}(X_{T_1} = j | X_0 = i). \end{aligned}$$

Ainsi

$$\begin{aligned} Q_{i,j} &= \lim_{t \rightarrow 0} \frac{P_{i,j}(t)}{t} \\ &= \lim_{t \rightarrow 0} \frac{\mathbb{P}(T_1 < t | X_0 = i)}{t} \mathbb{P}(X_{T_1} = j | X_0 = i) \end{aligned}$$

$$Q_{i,j} = \begin{cases} i\lambda\mu & \text{si } j = i - 1 \\ i\lambda(1 - \mu)\mathbb{P}(Y = k) & \text{si } j = i + k \end{cases}$$

■

Remarque : En utilisant le théorème (6) et les équations (2.3) et (2.4), nous pouvons facilement écrire les équations de Kolmogorov pour le processus X .

$$\frac{d}{dt} P_{i,j}(t) = \sum_{k=0}^{\infty} P_{i,k}(t) Q_{k,j} = -jaP_{i,j} + a \sum_{1 \leq k \leq j+1, k \neq i} kp_{j-k+1} P_{i,k}(t) \quad (\text{forward}), \quad (2.7)$$

$$\frac{d}{dt} P_{i,j}(t) = \sum_{k=0}^{\infty} Q_{i,k} P_{k,j}(t) = -jaP_{i,j} + ia \sum_{k \geq i-1, k \neq i} kp_{k-i+1} P_{k,j}(t) \quad (\text{backward}), \quad (2.8)$$

avec les conditions initiales

$$P_{i,j}(0+) = \begin{cases} 1 & \text{pour } i = j \\ 0 & \text{pour } i \neq j \end{cases}$$

Théorème 8. Soit l'égalité $V_j(z) = \partial_t G_j(0, z)$ qui est en relation avec le générateur infinitésimal Q comme suit :

$$V_j(z) = \sum_l Q_{j,l} z^l, \quad Q_{j,j} = -\sum_{l \neq k} Q_{k,l} = -\lambda_k.$$

Alors, de (2.5), on en déduit que :

$$\partial_t G(t, z) = V(z) \frac{\partial G(t, z)}{\partial z}. \quad (2.9)$$

Preuve.

$$\begin{aligned} \partial_t G(t, z) &= \lim_{s \rightarrow 0} \frac{G(t+s, z) - G(t, z)}{s} \\ &= \lim_{s \rightarrow 0} \frac{G(t, G(s, z)) - G(t, z)}{s} \\ &= \lim_{s \rightarrow 0} \frac{h(s) - h(0)}{s} = h'(0), \end{aligned}$$

avec $h(s) = G(t, G(s, z))$. Puisque $G(0, z) = z$, on a $h(0) = G(t, z)$, puis

$$h'(s) = \partial_z G(t, G(s, z)) \partial_s G(s, z).$$

D'où

$$\partial_t G(t, z) = h'(0) = \partial_z G(t, z) \partial_t G(0, z).$$

■

On cherche à trouver l'intégrale première du système qui vérifie l'EDO suivante :

$$\begin{cases} \partial_t \mathcal{I} = -V(\mathcal{I}) \\ \mathcal{I}(0, y) = y \end{cases}$$

dont la solution sera notée $\mathcal{I}(t, y)$

On sait que l'intégrale première $\mathcal{I}(t, y)$ est constante sur la courbe intégrale. Ainsi, si on veut avoir sa valeur au temps t quelconque, il suffit juste de connaître sa valeur à l'origine. Cela signifie que :

Théorème 9. Si $V(z)$ existe et différentiable, alors la solution de (2.9), $(t, z) \longrightarrow (t, G(t, z))$ est la réciproque $(t, y) \longrightarrow (t, \mathcal{I}(t, y))$.

Preuve. Soit

$$\begin{aligned} \begin{pmatrix} t \\ y \end{pmatrix} &\xrightarrow{A} \begin{pmatrix} t \\ \mathcal{I}(t, y) \end{pmatrix} \\ \begin{pmatrix} t \\ G(t, z) \end{pmatrix} &\xleftarrow{A^{-1}} \begin{pmatrix} t \\ z \end{pmatrix} \end{aligned}$$

leur Jacobien sont :

$$J_A(t, y) = \begin{pmatrix} 1 & 0 \\ \partial_t \mathcal{I} & \partial_y \mathcal{I} \end{pmatrix} \quad \text{et} \quad J_{A^{-1}}(t, z) = \begin{pmatrix} 1 & 0 \\ \partial_t G & \partial_z G \end{pmatrix}.$$

L'inverse de la matrice $J_A(t, y)$ vérifie $(J_A(t, y))^{-1} = J_{A^{-1}}(t, z)$ avec $z = \mathcal{I}(t, y)$. Ainsi,

$$(J_A(t, y))^{-1} = \begin{pmatrix} 1 & 0 \\ -\frac{\partial_t \mathcal{I}}{\partial_y \mathcal{I}(t, y)} & \frac{1}{\partial_y \mathcal{I}(t, y)} \end{pmatrix}$$

Donc on a :

$$\partial_t G(t, z) = -\frac{\partial_t \mathcal{I}(t, y)}{\partial_y \mathcal{I}(t, y)} = \frac{1}{\partial_y \mathcal{I}(t, y)} (V(\mathcal{I}(t, y)))$$

et

$$\partial_z G(t, z) = -\frac{1}{\partial_y \mathcal{I}(t, y)}.$$

Cela implique que $\partial_t G(t, z) = V(z) \partial_z G(t, z)$.

■

2.2.3 Probabilité d'extinction et moments

Soit $G(t, z)$ la fonction génératrice de $X(t)$, sachant que $X(0) = 1$. Elle s'écrit

$$G^{[1]}(t, z) = \sum_{k=0}^{\infty} p_k(t) z^k.$$

Ainsi $P(X(t) = 0) = G^{[1]}(0, t)$.

Soit $t_1 < t_2$. Comme l'événement $A = "X(t_1) = 0"$ implique l'événement $B = "X(t_2) = 0"$, on obtient $A \subset B$, puis

$$P(X(t_2) = 0) \geq P(X(t_1) = 0) \quad \text{ou encore} \quad G(t_2, 0) \geq G(t_1, 0),$$

Alors, $G(t, 0)$ est une fonction croissante en t , bornée et donc convergente. Notons

$$q = \lim_{t \rightarrow \infty} G(t, 0) \leq 1.$$

En utilisant l'équation (2.5), on a :

$$G(t + s, 0) = G(t, G(s, 0)) \Rightarrow q_{t+s} = G(t, q_s)$$

en faisant tendre s vers l'infini, on a $G(t, q_\infty) = q_\infty$, pour tout t .

Les fonctions génératrices sont convexes sur $[0, 1]$ (sauf pour la loi binomiale), car

$$g''(z) = \sum_{k \geq 1} k(k-1)z^k \geq 0.$$

Ainsi l'équation $G(t, z) = z$ ne peut avoir qu'une seule solution avec $z < 1$, car $G(t, 1) = 1$.

Nous avons alors

$$\begin{aligned} \partial_t G^{[1]}(t, z) &= \sum_{j \geq 0} \partial_t P_{1,j}(t) z^j \\ &= \sum_{j \geq 0} z^j \sum_{k \geq 0} Q_{1,k} P^{k,j}(t) \\ &= \sum_{k \geq 0} Q_{1,k} \sum_{j \geq 0} z^j P^{k,j}(t) \\ &= \sum_{k \geq 0} Q_{1,k} (G^{[1]}(t, z))^k, \end{aligned}$$

en usant la propriété de branchements indépendants des k individus initiaux. Or

$$Q_{1,1} = -\lambda \quad \text{et} \quad Q_{1,k} = \lambda p_k \quad \text{si} \quad k \neq 1.$$

D'où

$$\partial_t G^{[1]}(t, z) = \lambda \left(\sum_{k \geq 0} p_k (G^{[1]}(t, z))^k - G^{[1]}(t, z) \right) = \lambda (g(G^{[1]}(t, z)) - G^{[1]}(t, z))$$

où $g(z) = \sum_{k \neq 0} p_k z^k$ est la fonction génératrice de la loi de reproduction de Y conditionnée à $Y \neq 1$.

Si on pose $s = G^{[1]}(t, z)$ et $u(s) = \partial_t G^{[1]}(t, z)$, nous pouvons écrire

$$u(s) = \lambda (g(s) - s), \quad s \in [0, 1], \quad (2.10)$$

autrement dit

$$\partial_t G^{[1]}(t, z) = u(G^{[1]}(t, z)), \quad \forall z \in \mathcal{B}(0, 1), \quad (2.11)$$

avec les conditions frontières $\begin{cases} G^{[1]}(0, z) = z \\ G^{[1]}(t, 1) = 1 \end{cases}$

Équations des moments :

Les moments du processus, quand ils existent, peuvent également être obtenus implicitement à partir des équations de Kolmogorov, par exemple pour la moyenne. De (2.10), nous pouvons aussi déduire une équation pour les moments.

Soit $m_1(t) = E(X(t))$ et supposons que $m_1(t) < \infty$, on a alors

$$m_1(t) = \partial_z G^{[1]}(t, z) \Big|_{z=1}$$

d'où

$$\partial_t m_1(t) = \partial_z \partial_t G^{[1]}(t, z) \Big|_{z=1} = \partial_z u(G^{[1]}(t, z)) \Big|_{z=1}.$$

Ainsi $m_1(t)$ est une solution de l'équation différentielle

$$m_1'(t) = u'(G^{[1]}(t, 1)) m_1(t) = u'(1) m_1(t),$$

avec $u'(1) = \lambda(g'(1) - 1) = \lambda(m - 1)$ et $m = \sum_{k \neq 1} kp_k$.

Nous on déduisons que :

$$m_1(t) = e^{\lambda(m-1)t} \quad (2.12)$$

Remarquons que si $X(0) = k$, alors

$$m^{[k]}(t) = k m(t) = m(t)\mathbb{E}(X(0)).$$

En fonction de la moyenne m de la distribution des descendants conditionnellement à un événement, le processus de branchement se comporte différemment.

Ainsi, si $m > 1$, le processus est dit sur-critique et $m(t) \xrightarrow[t \rightarrow \infty]{} +\infty$. Les cas $m = 1$ et $m < 1$ correspondent à des processus critiques ($m(t) \equiv 1$) et sous-critiques ($m(t) \xrightarrow[t \rightarrow \infty]{} 0$). Dans ces deux derniers cas, on peut prouver en plus que le processus finira par disparaître presque sûrement.

Variance :

Supposons que le moment de second d'ordre, $m_2(t) = \mathbb{E}(X^2(t)) < \infty$.

Alors

$$\partial_{z^2}^2 G^{[1]}(t, z) |_{z=1} = m_2(t) - m_1(t),$$

$$\partial_t m_2(t) - \partial_t m_1(t) = \partial_t \partial_{z^2}^2 G^{[1]}(t, z) |_{z=1} = \partial_{z^2}^2 \partial_t G^{[1]}(t, z) |_{z=1} = \partial_{z^2}^2 u(G^{[1]}(t, z)) |_{z=1}.$$

En utilisant l'équation de Kolmogorov backward, $m_2(t)$ satisfait l'équation différentielle suivante :

$$m_2'(t) - m_1'(t) = u'(1)(m_2(t) - m_1(t)) + u''(1)(m_1(t))^2,$$

avec la condition $m_2(0) = 1$ et $u''(1) = \lambda g''(1) = \lambda(\sigma^2 - m + m^2)$, $u'(1) = \lambda(m - 1)$.

On a alors

$$m_1(t) = e^{u'(1)t}.$$

Au final, on obtient l'EDO linéaire non homogène :

$$m_2'(t) = u'(1)m_2(t) + u''(1)(m_1(t))^2.$$

La solution de ce qui précède est :

$$m_2(t) = e^{u'(1)t} m_2(0) + u''(1) e^{u'(1)t} \int_0^t e^{-u'(1)s} (m_1(s))^2 ds,$$

autrement dit

$$m_2(t) = e^{\lambda(m-1)t} \left(\theta e^{\lambda(m-1)t} - (\theta - 1) \right)$$

avec $\theta = \frac{\sigma^2}{m-1} + m$ si $m \neq 1$ et $m_2(t) = \lambda \sigma^2 t$ si $m = 1$.

Donc la variance s'écrit sous la forme suivante

$$Var(t) = m_2(t) - m_1^2(t) = (\theta - 1) e^{\lambda(m-1)t} \left(e^{\lambda(m-1)t} - 1 \right), \text{ si } m \neq 1 \quad (2.13)$$

et

$$Var(t) = \lambda \sigma^2 t - 1, \quad \text{si } m = 1. \quad (2.14)$$

Définition 8. Le processus $X(t)$ est dit régulier si $X(t) < \infty$ presque sûrement pour chaque $t > 0$. Dans le cas contraire, il est dit non régulier. En particulier, $X(t)$ est régulier dès que $X(t)$ est critique ou sous-critique.

dire qu'un processus est non-régulier veut donc dire qu'il existe $t_* < \infty$ tel que

$$\mathbb{P}(X(t_*) = \infty) > 0.$$

2.3 Quelques exemples simples de branchement

Développons maintenant quelques exemples pour illustrer la diversité de la nature des équations de branchement.

2.3.1 Loi de Bernoulli pour la descendance

Ici, on considère que Y suit la loi de Bernoulli

$$\begin{cases} \mathbb{P}(Y = 0 \mid Y \neq 1) = \mu \\ \mathbb{P}(Y = 2 \mid Y \neq 1) = 1 - \mu \end{cases}$$

Ainsi on a $m = 2(1 - \mu)$, par conséquent, si $\mu < \frac{1}{2}$ le processus est sur-critique et pour $\mu \geq \frac{1}{2}$ on aura la situation de processus critique ou sous-critique. On a de plus que :

$$\text{Var}(Y \mid Y \neq 1) = 4\mu(1 - \mu).$$

Dans le cas $\mu \neq \frac{1}{2}$, de (2.13), on déduit

$$\begin{cases} m(t) = e^{\lambda(1-2\mu)t} \\ \text{Var}(t) = \frac{e^{\lambda(1-2\mu)t}}{1-2\mu} (e^{\lambda(1-2\mu)t} - 1) \end{cases}$$

et dans le cas critique où $\mu = \frac{1}{2}$, de (2.14), on a :

$$\begin{cases} m(t) = 1 \\ \text{Var}(t) = \lambda 4\mu(1 - \mu)t \end{cases}$$

Son générateur infinitésimal sera défini par :

$$Q_{1,j} = \begin{cases} \lambda\mu & \text{si } j = 0 \\ -\lambda & \text{si } j = 1 \\ \lambda(1 - \mu) & \text{si } j = 2 \\ 0 & \text{si } j > 2 \end{cases}$$

Nous cherchons la solution générale de l'équation linéaire aux dérivées partielles du branchement :

$$\begin{cases} \partial_t G(t, z) = V(z) \partial_z G(t, z) \\ G(0, z) = z. \end{cases} \quad (2.15)$$

Pour cela on cherche à trouver l'intégrale première du système qui vérifie l'EDO suivant :

$$\begin{cases} \partial_t \mathcal{I} = -V(\mathcal{I}) \\ \mathcal{I}(0, z) = z \end{cases}.$$

Sa solution sera notée $\mathcal{I}(t, z)$

Écrivons les équations de branchement. On a ici :

$$\begin{aligned} V(z) &= \sum_{k \geq 0} Q_{1,j} z^j \\ &= \lambda\mu - \lambda z + \lambda(1 - \mu)z^2. \end{aligned}$$

Donc on peut écrire l'EDO suivante :

$$\partial_t \mathcal{I} = -V(\mathcal{I}) = -(\lambda\mu - \lambda\mathcal{I} + \lambda(1 - \mu)\mathcal{I}^2).$$

Si on pose

$$A = -\lambda, \quad B = -\lambda(1 - \mu), \quad C = -\lambda\mu$$

L'équation devient $\mathcal{I}'(t, z) + A\mathcal{I}(t, z) = B\mathcal{I}^2(t, z) + C$, cette simplification nous conduit à l'équation différentielle dite de Riccati.

- L'intégration d'une équation différentielle de Riccati nécessite la connaissance d'une solution particulière θ de cette équation, telle que $\theta = A \pm \sqrt{A^2 - 4BC}$, autrement dit $\theta = \lambda(-1 \pm \sqrt{1 - 4\mu(1 - \mu)})$.
- Le changement de fonction inconnue $y = \mathcal{I}(t) - \theta$ transforme l'équation différentielle :

$$\mathcal{I}'(t, z) + A\mathcal{I}(t) = B\mathcal{I}^2(t, z) + C$$

en

$$(y + \theta)' + A(y + \theta) = B(y + \theta)^2 + C, \text{ ce qui se simplifie en :}$$

$$y' + Ay + \theta' + A\theta = By^2 + 2B\theta y + B\theta^2 + C.$$

Or une solution particulière θ de l'équation de Riccati vérifie : $\theta' + A\theta = B\theta^2 + C$. Cette simplification nous conduit à l'équation différentielle de **Bernoulli** ($\alpha = 2$).

$$y'(A - 2B\theta)y = By^2.$$

On posera un nouveau changement de fonction $u = \frac{1}{y}$ pour se ramener à une équation différentielle linéaire en u : $-u' + [A - 2B\theta]u = B$.

La solution de cette équation est

$$u = e^{(A-2B\theta)t} \left[\frac{B}{2B\theta - A} e^{-(A-2B\theta)t} + K \right]$$

Après résolution de cette dernière équation, la solution $\mathcal{I}(t)$ de l'équation de Riccati sera obtenue par : $\mathcal{I}(t, z) = y(t) + \theta$ avec

$$y(t) = \frac{1}{u} \quad (\text{où } \theta = A \pm \sqrt{A^2 - 4BC}).$$

$$\text{On pose } \begin{cases} \alpha = A - 2B\theta \\ \beta = \frac{B}{2B\theta - A} \end{cases} \Rightarrow \begin{cases} \alpha = \lambda(2\theta(1 - \mu) - 1) \\ \beta = \frac{\mu - 1}{1 - 2(1 - \mu)\theta} \end{cases}, \text{ ce qui donne :}$$

$$\mathcal{I}(t, z) = \frac{e^{-\alpha t}}{\beta e^{-\alpha t} + K} + \theta.$$

Or, on a $\mathcal{I}(0, z) = z = \frac{1}{\beta + K} + \theta$. Donc d'où $K = \frac{1 - \beta z + \beta \theta}{z - \theta}$ et la solution $\mathcal{I}(t, z)$ est donnée par

$$G(t, z) = \frac{e^{-\alpha t}(z - \theta)}{1 + \beta(z - \theta)(e^{-\alpha t} - 1)} + \theta.$$

Autrement dit, si on pose $\gamma = 1 - 2(1 - \mu)\theta$, on a

$$G(t, z) = \frac{(z - \theta)(2\mu\theta - 1)\gamma e^{\lambda\gamma t}}{\gamma + (\mu - 1)(z - \theta)(e^{\lambda\gamma t} - 1)} + \theta.$$

2.3.2 Loi géométrique pour la descendance

Si Y suit la loi géométrique de paramètre φ ,

$$\mathbb{P}(Y = k) = (1 - \varphi)^{k-1} \varphi.$$

Ainsi on a $m = \frac{1}{p}$, et $\sigma^2 = \frac{1-p}{p^2}$, par conséquent si $p \leq 1$ on est dans le cas critique ou sous-critique (l'extinction est certaine) et si $p > 1$ sur-critique (la probabilité survivra ou explosera).

Dans le cas $p \neq 1$ de (2.13), on a :

$$\begin{cases} m(t) = e^{\lambda(\frac{1-p}{p})t} \\ Var(t) = \frac{(2-p)e^{\lambda(\frac{1-p}{p})t}}{p} \left(e^{\lambda(\frac{1-p}{p})t} - 1 \right) \end{cases}$$

et dans le cas critique où $p = 1$, de (2.14), on déduit

$$\begin{cases} m(t) = 1 \\ Var(t) = \frac{\lambda(1-p)t}{p^2} \end{cases}.$$

Calculons maintenant le générateur infinitésimal, défini par

$$Q_{1,j} = \begin{cases} \lambda\mu & \text{si } j = 0 \\ -\lambda & \text{si } j = 1 \\ \lambda(1 - \mu)P(y = k) & \text{si } j = k + 1 \end{cases}$$

On sait que

$$\begin{aligned}
 V(z) &= \sum_{k \geq 0} Q_{1,k} z^k \\
 &= \lambda\mu - \lambda z + \lambda(1 - \mu) \sum_{k \geq 1} P(y = k) z^{k+1} \\
 &= \lambda\mu - \lambda z + \frac{\lambda\varphi(1 - \mu)z}{1 - \varphi} \left(\sum_{k \geq 0} ((1 - \varphi)z)^k - 1 \right) \\
 &= \lambda\mu - \lambda z + \frac{\lambda\varphi(1 - \mu)z}{1 - \varphi} \left(\frac{1}{1 - (1 - \varphi)z} - 1 \right) \\
 &= \lambda\mu - \lambda z + \frac{\lambda\varphi(1 - \mu)z^2}{1 - (1 - \varphi)z} \\
 &= \frac{\lambda\mu - \lambda z(1 + \mu(1 - \varphi)) + \lambda z^2(1 - \mu\varphi)}{1 - (1 - \varphi)z}.
 \end{aligned}$$

Pour écrire des equations de branchement , on cherche a trouver l'intégrale première du système qui vérifie l'EDO suivante :

$$\begin{cases} \partial_t \mathcal{I} = -V(\mathcal{I}) \\ \mathcal{I}(0, y) = y \end{cases}.$$

D'où

$$\partial_t \mathcal{I} = -V(\mathcal{I}) = \frac{-\lambda\mu + \lambda\mathcal{I}(1 + \mu(1 - \varphi)) - \lambda\mathcal{I}^2(1 - \mu\varphi)}{1 - (1 - \varphi)\mathcal{I}}.$$

On pose :

$$\begin{cases} A = -\lambda(1 - \mu\varphi), & B = \lambda(1 + \mu(1 - \varphi)) \\ C = -\lambda\mu, & D = 1, \quad E = -(1 - \varphi). \end{cases}$$

On obtient :

$$\mathcal{I}'(t, y) = \frac{A\mathcal{I}^2(t, y) + B\mathcal{I}(t, y) + C}{D + E\mathcal{I}(t, y)},$$

autrement dit

$$\frac{\mathcal{I}'(t, y)(D + E\mathcal{I}(t, y))}{A\mathcal{I}^2(t, y) + B\mathcal{I}(t, y) + C} = 1,$$

d'où

$$\int_a^t \left(\frac{\mathcal{I}'(t, y)(D + E\mathcal{I}(t, y))}{A\mathcal{I}^2(t, y) + B\mathcal{I}(t, y) + C} \right) dt = t - a,$$

c.à.d.

$$\frac{\mathcal{I}'(t, y)(D + E\mathcal{I}(t, y))}{A\mathcal{I}^2(t, y) + B\mathcal{I}(t, y) + C} = \frac{\alpha}{(\mathcal{I}(t, y) - \mathcal{I}_1)} + \frac{\beta}{\mathcal{I}(t, y) - \mathcal{I}_2}$$

tel que

$$\begin{cases} \alpha &= \frac{-\mathcal{I}_1 E + D}{\mathcal{I}_2 - \mathcal{I}_1}, & \beta &= \frac{\mathcal{I}_2 E - D}{\mathcal{I}_2 - \mathcal{I}_1}, \\ \mathcal{I}_1, \mathcal{I}_2 &= \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}. \end{cases}$$

$$\text{D'où } \int \left(\frac{D + E\mathcal{I}(t, y)}{A\mathcal{I}^2(t, y) + B\mathcal{I}(t, y) + C} \right) dt = t - a$$

$$\alpha \int \left(\frac{1}{\mathcal{I}(t, y) - \mathcal{I}_1} \right) dt + \beta \int \left(\frac{1}{\mathcal{I}(t, y) - \mathcal{I}_2} \right) dt = t - a$$

$$\ln(\mathcal{I}(t, y) - \mathcal{I}_1)^\alpha (\mathcal{I}(t, y) - \mathcal{I}_2)^\beta = t - a.$$

Donc $\mathcal{I}(t, y)$ est la solution de l'équation implicite

$$(\mathcal{I}(t, y) - \mathcal{I}_1)^\alpha (\mathcal{I}(t, y) - \mathcal{I}_2)^\beta = e^{t-a}$$

D'après le théorème (8), $(t, G(t, z))$ est la réciproque de $(t, \mathcal{I}(t, y))$.

2.3.3 Loi de Poisson

Si Y suit la loi de Poisson de paramètre ρ , $\mathbb{P}(Y = k) = \frac{\rho^k}{k!} e^{-\rho}$.

On a $m = \sigma^2 = \rho$. Par conséquent si $\rho \leq 1$, le processus est critique ou sous-critique (l'extinction est certaine) et si $\rho > 1$ il est sur-critique (probabilité de survivre et d'exploser).

Dans le cas $\rho \neq 1$, de (2.13) on déduit

$$\begin{cases} m(t) = e^{\lambda(\rho-1)t} \\ Var(t) = (\rho + \frac{1}{\rho-1})e^{\lambda(\rho-1)t} (e^{\lambda(\rho-1)t} - 1) \end{cases}$$

et dans le cas critique où $\rho = 1$, de (2.14), on a

$$\begin{cases} m(t) = 1 \\ Var(t) = \lambda \rho t \end{cases}$$

Par ailleurs, nous avons les relations suivantes :

$$\begin{aligned} V(z) &= \sum_{j \geq 0} Q_{1,j} z^j \\ &= \lambda \mu - \lambda z + \lambda \gamma \sum_{k \geq 1} P(y = k) z^{k+1} \\ &= \mu \lambda - \lambda z - \gamma \lambda e^{-\rho} z \left(\sum_{k \geq 1} \frac{(\rho z)^k}{k!} \right) \\ &= \mu \lambda - \lambda z - \gamma \lambda e^{-\rho} z (e^{\rho z} - 1) \end{aligned}$$

D'ici, on va chercher à trouver l'intégrale première du système qui vérifie l'EDO suivante :

$$\begin{cases} \partial_t \mathcal{I} = -V(\mathcal{I}) \\ \mathcal{I}(0, y) = y \end{cases}.$$

Donc

$$\mathcal{I}'(t, y) = -\mu \lambda + \lambda \mathcal{I}(t, y) + \gamma \lambda e^{-\rho} \mathcal{I}(t, y) (e^{\rho \mathcal{I}(t, y)} - 1).$$

On n'a pas de formule achevée mais une relation qui nécessite des approximations numériques.

2.4 Processus de branchement multi-types homogènes en temps continu

Tout au long de cette thèse, nous utiliserons les notations suivantes :

K désignera le nombre de sous populations ou classes, $\mathbb{N} := \{0, 1, 2, \dots\}$, $\mathbb{N}^* := \{1, 2, \dots\}$, $\mathbb{N}_+^K := [0, \infty[^K$, $\mathbb{R}_+ := [0, \infty[$.

Plus particulièrement, on distinguera les points ou états suivants :

$$\mathbf{0} = (0, \dots, 0), \quad \mathbf{1} = (1, \dots, 1) \in \mathbb{N}^K,$$

$\mathbf{e}_i = (0, \dots, 0, \underset{i}{1}, 0, \dots, 0)$, qui est l'état de la population qui contient un seul individu de classe i .

De façon plus générale, $i = (i_1, \dots, i_K)$, $j = (j_1, \dots, j_K) \in \mathbb{N}^K$ sont des multi-indices décrivant l'état des K sous populations.

On considérera aussi des variables complexes $\mathbf{z} = (z_1, \dots, z_K) \in D^K(0, 1)$ où $D^K(0, 1)$ est le disque ouvert dans \mathbb{C}^K , et on notera :

$$\mathbf{z}^i = (z_1^{i_1}, \dots, z_K^{i_K}) = \prod_{l=1}^K z_l^{i_l} \text{ et } (\mathbf{z}^i)_{[-j]} = \prod_{l \neq j} z_l^{i_l}.$$

On ne distinguera pas sauf si c'est nécessaire, entre les indices uni ou K -dimensionnels.

Le processus de branchement multi-types classique est un processus de Markov homogène $\mathbf{X}(t) = (X_1(t), \dots, X_K(t))$, dont les états sont des vecteurs de \mathbb{N}^K : $\mathbf{X}(t)$ décrit le nombre d'individus de chaque type à l'instant t . Chaque individu d'une classe k vit un temps exponentiel de paramètre λ_k et donne naissance à plusieurs individus répartis dans les différentes classes.

Définition 9. *Un processus de branchement multi-types est donné par K populations évoluant en interaction, et où chaque individu de la population k*

- *évolue indépendamment de tous les autres et de sa date de naissance,*
- *au bout d'un laps de temps aléatoire (de loi propre à sa classe), cet individu, meurt, ou se reproduit en plusieurs individus répartis dans les différentes sous populations selon une loi propre à sa classe,*
- *les descendants se rajoutent aux classes correspondantes.*

2.4.1 Outil de base : encore la fonction génératrice

Pour les processus de branchement multi-types, les fonctions génératrices vont jouer le même rôle fondamental que dans le cas uni-type.

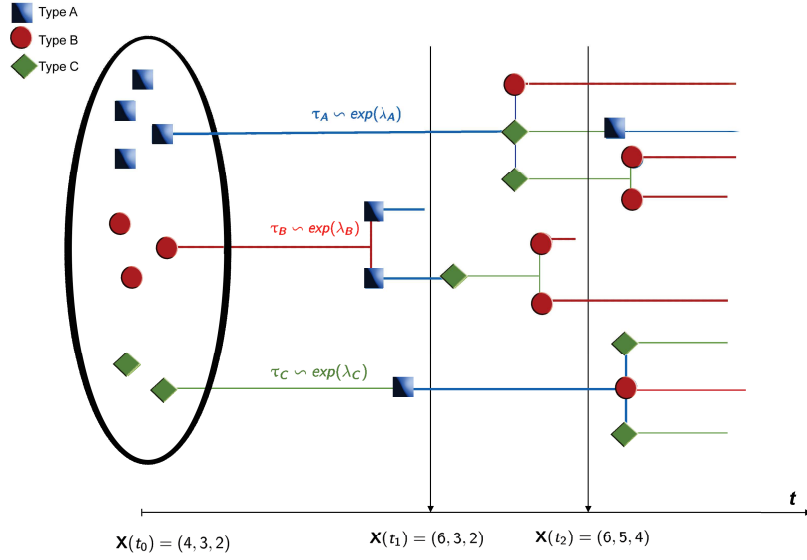


FIGURE 2.1 – Branchement multi-types d'ordre 3.

Soit $\mathbf{z} = (z_1, z_2, \dots, z_K) \in D^K(0, 1)$, on considère les fonctions génératrices de $\mathbf{X}(t)$ conditionnellement à $\mathbf{X}(0) = e_i$,

$$\begin{aligned} G_{e_i}(t, \mathbf{z}) &= E(z_1^{X_1(t)} \times \dots \times z_K^{X_K(t)} \mid X(0) = e_i) \\ &= E(\mathbf{z}^{\mathbf{X}(t)} \mid \mathbf{X}(0) = e_i). \end{aligned}$$

Ainsi,

$$G_{e_i}(t, \mathbf{z}) = \sum_{\mathbf{l} \in \mathbb{N}^K} P_{i,1}(t) \mathbf{z}^{\mathbf{l}}$$

Pour simplifier la présentation, on note

$$G_i(t, \mathbf{z}) = G_{e_i}(t, \mathbf{z}) \text{ et } \mathbf{G}(t, \mathbf{z}) = (G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})).$$

Ainsi on trouve la fonction génératrice qui déterminera la distribution du nombre de descendants vivants de chaque type à l'instant t :

- Si $\mathbf{X}(0) = (h_1, \dots, h_K)$ alors $\mathbf{G}_{(h_1, \dots, h_K)}(t, \mathbf{z}) = G_1^{h_1}(t, \mathbf{z}) \times \dots \times G_K^{h_K}(t, \mathbf{z})$.
- Si $\mathbf{X}(0)$ est de loi γ_0 , on utilisera l'extension, $\mathbf{G}_{\gamma_0}(t, \mathbf{z}) = \mathbf{G}_{\gamma_0}(\mathbf{G}(t, \mathbf{z}))$.

2.4.2 Équation de Kolmogorov

On déduit, à partir des hypothèses de branchement, l'équation vectorielle suivante, telle que pour tous $s, t \geq 0$ et $\mathbf{z} \in D^K(0, 1)^K$ on a

$$\mathbf{G}(t + s, \mathbf{z}) = \mathbf{G}(t, \mathbf{G}(s, \mathbf{z})) \quad \text{avec} \quad \mathbf{G}(0, \mathbf{z}) = \mathbf{z}.$$

La propriété fondamentale de l'indépendance des branchements se traduit de la façon suivante : Si $\mathbf{X}(0) = \mathbf{e}_i$, on a

$$G_i(t, \mathbf{z}) = \sum_{\mathbf{j}} P_{i,\mathbf{j}}(t) \mathbf{z}^{\mathbf{j}} = \prod_{m=1}^K [G_h(t, \mathbf{z})]^{i_h} = G^i(t, \mathbf{z}), \quad \mathbf{z} \in D^K(0, 1).$$

Comme $\mathbf{X}(t)$ est un processus markovien de sauts et homogène dans le temps, on peut de manière explicite calculer son générateur infinitésimal Q associé au semi groupe $(P(t))_{t \geq 0} = (P_{i,j}(t), i, j \in \mathbb{N}^K)$ par

$$Q_{i,i} = \lim_{t \rightarrow 0} \frac{P_{i,i}(t) - 1}{t} = \partial_t P_{i,i}(t),$$

$$Q_{i,j} = \lim_{t \rightarrow 0} \frac{P_{i,j}(t)}{t} = \partial_t P_{i,j}(0), \quad i \neq j.$$

De façon concrète, on en déduit que la probabilité de transition $(P_{i,j}(t), t \geq 0)$ vérifie :

$$\partial_t P_{i,j}(t) = \sum_{h \geq 0} Q_{i,h} P_{h,i}(t),$$

i.e. $\partial_t P = QP$ en notation matricielle infini dimensionnelle.

De même, pour tout $h = (h_1, \dots, h_K)$, $\mathbf{z} \in D^K(0, 1)$, on a

$$\partial_t \mathbf{G}(t, \mathbf{z}) = \sum_l Q_{j,l} \mathbf{z}^l$$

et donc

$$\partial_t G_j(0, \mathbf{z}) = \sum_h Q_{j,h} \mathbf{z}^h$$

notons

$$V_j(\mathbf{z}) = \sum_h Q_{j,h} \mathbf{z}^h \text{ et } V(\mathbf{z}) = (V_1(\mathbf{z}), \dots, V_K(\mathbf{z}))$$

De façon similaire au cas uni-type, on en déduit dans le cas vectoriel :

$$\begin{cases} \partial_t \mathbf{G}(t, \mathbf{z}) &= V(\mathbf{z}) \partial_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) \\ \text{avec } V(\mathbf{z}) &= \partial_t \mathbf{G}(0, \mathbf{z}). \end{cases} \quad (2.16)$$

Par ailleurs, remarquons que chaque ligne de l'équation G_j ne met en jeu qu'un seul G_i (et les dérivées en $t = 0$ des autres G_i). C'est donc une équation aux dérivées partielles linéaire homogène d'une seule fonction.

C'est la même équation et seule la condition initiale change d'une ligne à l'autre. C'est une équation de type transport. Cela revient donc à résoudre l'équation :

$$\begin{cases} \partial_t \mathbf{G}(t, \mathbf{z}) = \sum_{j=1}^k V_j(\mathbf{z}) \partial_{z_j} \mathbf{G}(t, \mathbf{z}), \\ \mathbf{G}(0, \mathbf{z}) = \mathbf{z}. \end{cases} \quad (2.17)$$

avec

$$V_j(\mathbf{z}) = \partial_t G_j(0, \mathbf{z}).$$

2.4.3 Probabilité d'extinction et usage de la fonction génératrice des moments

De la même façon que pour PB uni-type, chaque $q_i(t) = P(\mathbf{X}(t) = \mathbf{0} \mid \mathbf{X}(0) = \mathbf{e}_i)$ est une fonction croissante de t , bornée donc converge. Soit $\mathbf{q} = \lim_{t \rightarrow \infty} \mathbf{q}(t)$, le vecteur de probabilité d'extinction, avec les conditions initiales distinctes $\mathbf{X}(0) = \mathbf{e}_i$ où

$$\mathbf{q}(t) = (q_1(t), \dots, q_K(t)), \quad q_i(t) = P_{\mathbf{e}_i}(\mathbf{X}_t = \mathbf{0}), \quad i = 1, \dots, K.$$

Par définition de la fonction génératrice, on obtient $\mathbf{q}(t) = \mathbf{G}(t, \mathbf{0})$. Par ailleurs, la propriété de branchement implique que pour tout $\mathbf{j} \in \mathbb{N}^K$,

$$\mathbb{P}_{\mathbf{j}}(\mathbf{X}_t = \mathbf{0}) = \mathbf{q}^{\mathbf{j}}(t) = q_1^{j_1}(t) \times \dots \times q_K^{j_K}(t),$$

ainsi,

$$\mathbb{P}_{\mathbf{j}} \left(\lim_{t \rightarrow \infty} \mathbf{X}_t = \mathbf{0} \right) = \mathbf{q}^{\mathbf{j}}.$$

Propriétés des fonctions génératrices des moments en multidimensionnel.

Soit X une v.a. sur \mathbb{N}^K , de loi $P = (p_h, h \in \mathbb{N}^K)$ et de fonction génératrice

$$g(\mathbf{z}) = E(\mathbf{z}^{\mathbf{X}}) = \sum_h p_h \mathbf{z}^h.$$

Comme en unidimensionnel, la fonction génératrice a des propriétés caractéristiques qui permettent de retrouver la loi de probabilité du moins de façon théorique mais aussi possède des propriétés analytiques et pratiques d'importance :

Ainsi,

— g est analytique par rapport à $z \in D^K(0, 1)$.

— On a $g(0) = p_0$, $g(1) = 1$ et plus généralement

$$p_h = \frac{\partial_{\mathbf{z}^h}^h g(0)}{h!}.$$

— Les marginales $\mathbf{X}_l(t)$ ont pour génératrices $g(1, \dots, 1, z_l, 1, \dots, 1) = g_l(z_l)$. Ainsi,

$$\mathbb{P}(\mathbf{X}_l = h_l) = \partial_{z_l}^{h_l} g(1, \dots, 1, 0, 1, \dots, 1).$$

— Si $l < l'$, les marginales bidimensionnelles $(X_l, X_{l'})$, ont pour génératrice

$$g_{ll'}(z_l, z_{l'}) = g(1, \dots, 1, z_l, 1, \dots, 1, z_{l'}, 1, \dots, 1) \text{ et sont finies.}$$

Concernant les moments, on a :

— Les moments d'ordre $r = (r_1, \dots, r_K)$ existent. Par conséquent, $\lim_{\mathbf{z} \rightarrow \mathbf{1}} \partial_{\mathbf{z}}^r g(\mathbf{z})$ existent.

Dans ce cas, on a :

$$\begin{aligned} \mathbb{E}(X_i) &= \partial_{z_i} g(1) \\ \mathbb{E}(X_i^2) &= \partial_{z_i}^2 g(1) + \partial_{z_i} g(1) \quad \text{si } i = j \\ \mathbb{E}(X_i X_j) &= \partial_{z_i z_j}^2 g(1) \quad \text{seulement si } i \neq j \end{aligned}$$

Ainsi, on a la matrice

$$\mathbb{E}(\mathbf{X}^T \mathbf{X}) = \left(\partial_{z_i, z_j}^2 g(1) \right) + \text{diag}(\mathbb{E}(X))$$

— Les lois conditionnelles ont des génératrices que l'on peut exprimer à partir de g (et de ses dérivées en 0).

Application au processus de branchement

On considère les vecteurs $m_i(t) = (\mathbb{E}(X(t) \mid X(0) = \mathbf{e}_i))$, $i = 1, \dots, K$
 et $w_l = (\partial_z V_l)(\mathbf{1})$.

Soient les fonctions matricielles $N^i(t) = \left(\partial_{z_j, z_{j'}}^2 G_i \right)(t, \mathbf{1})$ et $H_l = \left(\partial_{z_j, z_{j'}}^2 V_l \right)(1)$.

On remarque que $H_l = \left(\sum_h Q_{lh} h_j h_{j'} \right) - \text{diag}(w_l)$ correspond à la matrice des moments issus de la génératrice (instantanée) i.e. de la loi du nombre de descendants.

Proposition 4. *Soit la matrice $M_1(t) = E(\mathbf{X}(t)) = (m_{i,j}(t))$, $i, j = 1, \dots, K$ décrivant les nombres moyens d'individus de type j , conditionnellement aux états initiaux \mathbf{e}_i . Le moment d'ordre 1 s'écrit sous la forme suivante :*

$$M_1(t) = e^{tW} \tag{2.18}$$

où $W = (w_{lj})$ avec $w_{lj} = \partial_{z_j} V_l(\mathbf{1}) = \sum_h Q_{lh} h_j$ est la moyenne du nombre de descendants de type j issu du type l .

Pour tout i fixé, le moment d'ordre 2 s'écrit :

$$M_2(t) = (\mathbb{E}(X_j(t), X_{j'}(t) \mid \mathbf{X}(0) = \mathbf{e}_i))_{jj'} = N^i(t) + \text{diag}(m_i(t))$$

et vaut

$$M_2(t) = M_1^T(t) \left(\int_0^t (M_1^T)^{-1}(s) \Gamma(s) M_1(s) ds \right) M_1(t) \tag{2.19}$$

où $\Gamma(s) = \sum_i m_{il}(s) H_l$ et la covariance conditionnellement à $\mathbf{X}(0) = \mathbf{e}_i$ s'écrit :

$$\text{Cov}(\mathbf{X}(t)) = M_1^T(t) \left(\int_0^t (M_1^T)^{-1}(s) \Gamma(s) M_1(s) ds - I_K \right) M_1(t) \tag{2.20}$$

Preuve.

Soit la matrice $M_1(t) = (m_{i,j}(t))_{ij}$, avec

$$\begin{aligned} m_{i,j}(t) &= \mathbb{E}(X_j(t) \mid X(0) = \mathbf{e}_i) \\ &= \partial_{z_j} G_i(t, \mathbf{1}) \end{aligned}$$

Équation du vecteur moyenne : la relation $M_1(t) = E(\mathbf{X}(t)) = \partial_{\mathbf{z}}G(t, \mathbf{1})$ vérifie :

$$\begin{aligned}\partial_t m_{i,j}(t) &= \partial_t \partial_{z_j} G_i(t, \mathbf{z}) \Big|_{\mathbf{z}=\mathbf{1}} = \partial_{z_j} \partial_t G_i(t, \mathbf{z}) \Big|_{\mathbf{z}=\mathbf{1}} \\ &= \partial_{z_j} \left(\sum_{l=1}^K V_l(\mathbf{z}) \partial_{z_l} G_i(t, \mathbf{z}) \right) \Big|_{\mathbf{z}=\mathbf{1}} \\ &= \sum_{l=1}^K (\partial_{z_j} V_l(\mathbf{z}) \partial_{z_l} G_i(t, \mathbf{z}) + V_l(\mathbf{z}) \partial_{z_l, z_j}^2 G_i(t, \mathbf{z})) \Big|_{\mathbf{z}=\mathbf{1}}\end{aligned}$$

On a $V_l(\mathbf{z}) = \sum_h Q_{l,h} z^h$, avec

$$V_l(\mathbf{1}) = \mathbf{0}, \quad l = 1, \dots, K \text{ et } \partial_{z_j} V_l(\mathbf{z}) = \sum_h Q_{l,h} \partial_{z_j} (z_j^{h_j}) (\mathbf{z}^h)_{[-j]},$$

ce qui implique que $w_{lj} = \partial_{z_j} V_l(\mathbf{z}) \Big|_{\mathbf{z}=\mathbf{1}} = \sum_h Q_{l,h} h_j$.

Donc

$$\partial_t m_{i,j}(t) = \sum_{l=1}^K m_{i,l}(t) w_{lj} + 0.$$

Ainsi $M_1(t)$ est solution de l'équation différentielle :

$$\partial_t m_{i,j}(t) = \sum_{l=1}^K m_{i,l}(t) w_{lj}, \quad (2.21)$$

qui s'écrit au niveau matriciel, $\partial_t M_1(t) = M_1(t)W = WM_1(t)$. La solution de cette équation différentielle matricielle est

$$M_1(t) = M_1(0)e^{tW} = e^{tW},$$

car $M_1(0) = \text{diag}(\mathbf{1}) = I_K$.

Équation de la matrice de moments de second ordre

Elle vérifie :

$$M_2(t) = \mathbb{E}(X_j(t), X_{j'}(t) \mid \mathbf{X}(0) = \mathbf{e}_i) = \left(\partial_{z_j, z_{j'}}^2 G_i \right) (t, \mathbf{1}) + \text{diag}(m_i(t)),$$

Pour tout i fixé, on considère $m_i(t) = (m_{i,1}, \dots, m_{i,K})(t)$, et pour tous $j, j' = 1, \dots, K$

$$N_{jj'}^i(t) = \partial_{z_j, z_{j'}}^2 G_i(t, \mathbf{1}).$$

Alors, on a :

$$\begin{aligned}\partial_t N_{jj'}^i(t) &= \partial_t \partial_{z_j, z_{j'}}^2 G_i(t, \mathbf{1}) = \partial_{z_j, z_{j'}}^2 \partial_t G_i(t, \mathbf{1}) \\ &= \partial_{z_j, z_{j'}}^2 \left(\sum_{l=1}^K V_l(\mathbf{z}) \partial_{z_l} G_i(t, \mathbf{z}) \right) \Big|_{\mathbf{z}=\mathbf{1}} \\ &= \partial_{z_{j'}} \left(\sum_{l=1}^K (\partial_{z_j} V_l(\mathbf{z}) \partial_{z_l} G_i(t, \mathbf{z}) + V_l(\mathbf{z}) \partial_{z_l, z_j}^2 G_i(t, \mathbf{z})) \right) \Big|_{\mathbf{z}=\mathbf{1}} \\ &= \sum_{l=1}^K \left(\partial_{z_j, z_{j'}}^2 V_l(\mathbf{z}) \partial_{z_l} G_i(t, \mathbf{z}) + \partial_{z_j} V_l(\mathbf{z}) \partial_{z_l, z_{j'}} G_i(t, \mathbf{z}) \right. \\ &\quad \left. + \partial_{z_{j'}} V_l(\mathbf{z}) \partial_{z_l, z_j}^2 G_i(t, \mathbf{z}) + V_l(\mathbf{z}) \partial_{z_l, z_j, z_{j'}}^3 G_i(t, \mathbf{z}) \right) \Big|_{\mathbf{z}=\mathbf{1}}\end{aligned}$$

Puisque $V_l(\mathbf{1}) = \mathbf{0}$, pour tout $l = 1, \dots, K$, on obtient :

$$\partial_t N_{jj'}^i(t) = \sum_{l=1}^K \left(\partial_{z_j, z_{j'}}^2 V_l(\mathbf{1}) \partial_{z_l} G_i(t, \mathbf{1}) + \partial_{z_j} V_l(\mathbf{1}) \partial_{z_l, z_{j'}} G_i(t, \mathbf{1}) + \partial_{z_{j'}} V_l(\mathbf{1}) \partial_{z_l, z_j}^2 G_i(t, \mathbf{1}) + 0 \right)$$

On pose $W = (w_{lj})$ avec $w_{lj} = \partial_{z_j} V_l(\mathbf{z}) |_{\mathbf{z}=\mathbf{1}} = \sum_h Q_{lh} h_j$,

$$\begin{cases} H_{lj'} = \left(\partial_{z_j, z_{j'}}^2 V_l(\mathbf{1}) \right) = \left(\sum_h Q_{lh} h_j h_{j'} \right) \text{ si } j \neq j' \\ H_{ljj} = \left(\partial_{z_j}^2 V_l(\mathbf{1}) \right) = \left(\sum_h Q_{lh} h_j (h_j - 1) \right) = \left(\sum_h Q_{lh} h_j^2 - w_{lj} \right) \text{ si } j = j' \end{cases}$$

Ainsi :

$$\partial_t N_{jj'}^i(t) = \sum_{l=1}^K m_{il}(t) H_{h,j,j'} + \sum_{l=1}^K N_{j'l}^i(t) w_{lj} + \sum_{l=1}^K N_{jl}^i(t) w_{lj'}.$$

Remarquons que H_l et N^i sont des matrices symétriques, et qu'en posant

$$\Gamma_{jj'}(t) = \sum_{l=1}^K m_{il}(t) H_{lj'j'},$$

on obtient

$$\partial_t N_{jj'}^i(t) = \sum_{l=1}^K w_{lj} N_{lj'}^i(t) + \sum_{l=1}^K N_{jl}^i(t) w_{lj'} + \Gamma_{jj'}(t).$$

Au niveau matriciel, on aura l'équation différentielle linéaire non homogène suivante :

$$\partial_t N^i(t) = W^T N^i(t) + N^i(t) W + \Gamma(t). \quad (2.22)$$

Résolution de l'équation (2.22) dans le cas général

En s'inspirant de la résolution de l'équation de Riccati, considérons le système vectoriel de dimension K . Soit $(\tilde{X}(t), \tilde{Y}(t))$ une solution du système :

$$\tilde{X}' = A\tilde{X} + B(t)\tilde{Y} \quad (2.23)$$

$$\tilde{Y}' = C\tilde{Y} \quad (2.24)$$

où A , $B(t)$, C sont des matrices $(K \times K)$.

La solution de l'équation (2.24) s'écrit :

$$\tilde{Y}(t) = e^{Ct} \tilde{Y}(0),$$

où $\tilde{Y}(0)$ est une matrice donnée, par exemple $\tilde{Y}(0) = I_K$

l'équation (2.23) devient $\tilde{X}' = A\tilde{X} + B(t)e^{Ct}$ dont la solution est :

$$\tilde{X}(t) = e^{At} \left(\tilde{X}(0) + \int_0^t e^{-As} B(s) e^{-Cs} ds \right). \quad (2.25)$$

On remarquera que $\tilde{Y}(t)$ est toujours une matrice inversible et que $Z(t) = \tilde{X}(t)\tilde{Y}^{-1}(t)$ est solution de l'équation différentielle linéaire, non homogène, matricielle

$$Z'(t) = B(t) + AZ(t) - Z(t)C, \quad (2.26)$$

car

$$\begin{aligned} Z'(t) &= \tilde{X}'(t)\tilde{Y}^{-1} + \tilde{X}(t)\tilde{Y}^{-1}(t)\tilde{Y}'(t)\tilde{Y}^{-1}(t) \\ &= (A\tilde{X}(t) + B(t)\tilde{Y}(t))\tilde{Y}^{-1}(t) - \tilde{X}(t)\tilde{Y}^{-1}(t) \\ &= A\tilde{X}(t)\tilde{Y}^{-1}(t) - \tilde{X}(t)\tilde{Y}^{-1}(t)C + B(t) \\ &= AZ(t) - Z(t)C + B(t). \end{aligned}$$

Ainsi

$$Z(t) = e^{At} \left(\tilde{X}(0) + \int_0^t e^{-As} B(s) e^{-Cs} ds \right) e^{-Ct}. \quad (2.27)$$

Application à l'équation différentielle matricielle (2.22)

Prenons $A = W^T$, $B(t) = \Gamma(t)$, $C = -W$. Par conséquent, on a

$$\tilde{Y}(t) = e^{-tW} \quad (2.28)$$

$$\tilde{X}(t) = e^{tW^T} \left(\tilde{X}(0) + \int_0^t e^{-sW^T} \Gamma(s) e^{sW} ds \right). \quad (2.29)$$

De plus, on aura : $\tilde{Y}^{-1}(t) = e^{Wt} = M_1(t)$ et la solution générale de l'équation (2.22) vaut donc $N^i(t) = \tilde{X}(t)M_1(t)$.

Autrement dit, on a

$$N^i(t) = e^{tW^T} \left(N^i(0) + \int_0^t e^{-sW^T} \Gamma(s) e^{sW} ds \right) M_1(t) \quad (2.30)$$

avec la condition initiale $N^i(0)$. D'où, le moment de second ordre conditionnellement à $\mathbf{X}(0) = \mathbf{e}_i$:

$$M_2(t) = N^i(t) + \text{diag}(m_i(t)) = e^{tW^T} \left(N^i(0) + \int_0^t e^{-sW^T} \Gamma(s) e^{sW} ds \right) M_1(t) + \text{diag}(m_i(t)).$$

Autrement dit :

$$M_2(t) = M_1^T(t) \left(N^i(0) + \int_0^t (M_1^T)^{-1}(s) \Gamma(s) M_1(s) ds \right) M_1(t). \quad (2.31)$$

Puisque $M_2(0) = \text{diag}(\mathbf{e}_i) = N^i(0) + \text{diag}(\mathbf{e}_i)$, avec $N^i(0) = 0$, on obtient :

$$M_2(t) = M_1^T(t) \left(\int_0^t (M_1^T)^{-1}(s) \Gamma(s) M_1(s) ds \right) M_1(t) \quad (2.32)$$

et la covariance conditionnellement à $\mathbf{X}(0) = \mathbf{e}_i$:

$$\text{Cov}(\mathbf{X}(t)) = M_1^T(t) \left(\int_0^t (M_1^T)^{-1}(s) \Gamma(s) M_1(s) ds - I_K \right) M_1(t). \quad (2.33)$$

■

2.4.4 EDP, champs de vecteurs et courbes intégrales

Dans cette section, on rappelle quelques notations autour des équations aux dérivées partielles (EDP), des définitions et des propriétés des courbes intégrales et des intégrales premières. Ces dernières constituent une approche classique pour résoudre une EDP linéaire à partir d'une équation différentielle vectorielle, et de ses conditions initiales.

Soit un champ de vecteurs $W = (W_1, \dots, W_n)$ dans $\Omega \subset \mathbb{R}^n$ de classe C^1 . Une courbe intégrale de W est une fonction $\gamma(t) = (\gamma_1, \dots, \gamma_n)(t)$ d'un intervalle de \mathbb{R} à valeur dans Ω , dont la tangente est parallèle à W , c'est à dire qu'il existe une fonction scalaire $h(t)$ telle que $\gamma'(t) = h(t)W(\gamma(t))$.

Les choix possibles de $h(t)$ permettent de décrire la forme de la solution générale.

Ainsi, après une reparamétrisation par $h(t)$ de la courbe $\gamma(t)$, on aura une courbe intégrale solution de l'EDO vectorielle

$$\frac{d\mathbf{X}}{dt} = W(\mathbf{X}(t)). \quad (2.34)$$

Idée de la preuve :

Supposons que $H(t)$ (primitive de $h(t)$) admette une réciproque $\theta(t)$. Ainsi $\theta(t)$ est dérivable :

$$\theta'(t) = \frac{1}{H'(\theta(t))} = \frac{1}{h(\theta(t))}.$$

Il suffit de poser $\mathbf{X}(t) = \gamma(\theta(t))$, et de voir que

$$\mathbf{X}'(t) = \theta'(t)\gamma'(\theta(t)) = \theta'(t)h(\theta(t))W(\mathbf{X}(t)).$$

Soit maintenant, une solution passant par y en $t = 0$, notée $X(t, y)$. Sous nos hypothèses, $X(., .)$ sera différentiable par rapport à t et y .

L'EDO (2.34) est aussi écrite formellement sous la forme

$$\frac{dX_1}{W_1(X)} = \frac{dX_2}{W_2(X)} = \dots = \frac{dX_n}{W_n(X)}. \quad (2.35)$$

EDP linéaires du 1^{ère} ordre, intégrales premières

Une fonction $\mathcal{I} : \Omega \rightarrow \mathbb{R}$ est dite intégrale première du champ W , si elle est solution de l'EDP linéaire homogène :

$$\sum_{i=1}^n \partial_{y_i} \mathcal{I}(y) W_i(y) = 0, \quad \forall y \in \Omega. \quad (2.36)$$

On note : $\nabla = (\partial_{y_1}, \dots, \partial_{y_n})$, avec $y = (y_1, \dots, y_n)$

L'EDO $\partial_t X(t, y) = W(X(t, y))$ est dite "système caractéristique" de l'EDP (2.36). Cela signifie qu'on a $\nabla \mathcal{I} \perp W(y)$, autrement dit, d'après d'équation (2.36), $W(y)$ est dans le plan tangent de l'hypersurface $\mathcal{I} = \text{Cste}$.

On définit aussi une courbe intégrale de \mathcal{I} , comme étant une fonction $t \mapsto Y(t)$ telle que $\mathcal{I}(Y(t)) = Cste$.

Par ailleurs, on dit que les intégrales premières $\mathcal{I}_1(\cdot), \dots, \mathcal{I}_{n-1}(\cdot)$ sont fonctionnellement indépendantes si $\nabla \mathcal{I}_1, \dots, \nabla \mathcal{I}_{n-1}$ sont linéairement indépendants en tout $y \in \Omega$. S'il en est ainsi, on prouve que toute intégrale première \mathcal{I} est de la forme $\mathcal{I} = F(\mathcal{I}_1, \dots, \mathcal{I}_{n-1})$ où F est une fonction quelconque de classe C^1 . De cette façon, on prouve que toute courbe intégrale de W de \mathcal{I} , peut être obtenue car l'intersection des hypersurfaces

$$\mathcal{I}_1 = c_1, \dots, \mathcal{I}_{n-1} = c_{n-1},$$

en faisant varier les constants c_1, \dots, c_{n-1} .

Remarque :

Pour un champs de vecteur $W(y)$ et $\mathcal{I}(y)$ une intégrale première de W , on aura

$$\sum_i^n \partial_{y_i} \mathcal{I}(y) W_i(y) = 0 \tag{2.37}$$

c.à.d. solution de l'EDP (2.36), et $\mathcal{I} = F(\mathcal{I}_1, \mathbf{q}; \dots, \mathcal{I}_{n-1})$ avec des $\mathcal{I}_1, \dots, \mathcal{I}_n$ fonctionnellement indépendantes.

Si on considère l'hypersurface $\mathcal{I}(y) = Cste = c$, et si par exemple $W_1(y) \neq 0$, on peut considérer par un changement de paramètre $W_1(x_1)$ comme paramètre x_1 et décrire les courbes intégrales comme fonction $x_1 \mapsto (X_2(x_1), \dots, X_n(x_1))$.

Posons $X(y_1) = (X_1(x_1), X_2(y_1), \dots, X_n(y_1))$ avec $Y_1(y_1) = y_1$, on a :

$$X'_j(x_1) = \frac{W_j}{W_1}(X(x_1)), \quad \forall j = 1, \dots, n, \tag{2.38}$$

car en posant $r(y_1) = \mathcal{I}(X(x_1))$ on obtient :

$$\begin{aligned} r'(x_1) &= \sum_{i=1}^n \partial_{x_1} X_j(x_1) \partial_{x_j} \mathcal{I}(X(x_1)) \\ &= \frac{W_1}{W_1} \partial_{x_1} \mathcal{I}(X(x_1)) + \sum_{i=2}^n \frac{W_j}{W_1} \partial_{x_j} \mathcal{I}(X(x_1)) \\ &= \frac{1}{W_1} \left(\sum_{i=1}^n W_j \partial_{y_i} \mathcal{I}(X(x_1)) \right) = 0. \end{aligned}$$

Rechercher des solutions de l' EDP (2.36) équivaut à rechercher les intégrales premières indépendantes.

Voici formellement comment on peut trouver les $(n - 1)$ intégrales premières indépendantes. Si $W_1 \neq 0$, on réécrit le système (2.35) comme une nouvelle EDO avec $X_1 = x_1$ comme variable temps :

$$\frac{dX_2}{dx_1} = \frac{W_2}{W_1}, \dots, \frac{dX_n}{dx_1} = \frac{W_n}{W_1},$$

On notera la solution passant par u_2, \dots, u_n quand $x_1 = 0$ par $\phi(x_1; u_2, \dots, u_n)$.

On note $\phi_{x_1} : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$ qui est de classe C^1 , définie par

$$\phi_{x_1}(u_2, \dots, u_n) = (\phi_2(x_1, u_2, \dots, u_n), \dots, \phi_n(x_1, u_2, \dots, u_n)).$$

Par l'unicité des solutions pour l'EDO, la jacobienne de ϕ_{x_1} est inversible.

La réciproque ψ_{x_1} de $\phi_{x_1} : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$ a pour matrice jacobienne l'inverse $\psi_{x_1} = (\psi_2, \dots, \psi_n)$ et $\psi_j(x_1, x_2, \dots, x_n)$ sont des intégrales premières de W fonctionnellement indépendantes.

Maintenant pour choisir les $(n-1)$ intégrales premières fonctionnellement indépendantes, il suffit de remarquer que à tout $x = (x_1, x_2, \dots, x_n)$ la fonction

$$\mathcal{I}_j(x_1, x_2, \dots, x_n) = \psi_j(x_1, X_2(x_1), \dots, X_n(x_1)), \mathbf{j} = 2, \dots, n,$$

correspond à la $j^{\text{ème}}$ coordonnée à l'origine de la courbe intégrale $X(x_1)$ passant par (x_1, \dots, x_n) à "l'instant x_1 " et qu'elles sont indépendantes.

2.4.5 Application à la résolution de l'EDP "Branchement"

On a vu qu'en fait les équations (2.17) se résument à la résolution d'une seule EDP d'une fonction scalaire $g(t, z)$ avec des conditions initiales distinctes. Dans notre cas, $n = K + 1$ (t correspond à la variable X_1 dans (2.36)) et $W = (1, -V_1, -V_2, \dots, -V_K)$.

Ainsi, W ne s'annule jamais et le changement de variable $W_1(X) \longleftrightarrow x_1$ est ici déjà réalisé. On a :

$$\partial_t g(t, \mathbf{z}) = \sum_{m=1}^K V_m(\mathbf{z}) \partial_{z_m} g(t, \mathbf{z}), \quad (2.39)$$

où les V_m sont des fonctions analytiques sur les $B^K(0, 1)$ dans \mathbb{C}^K avec la condition initiale

$$g(0, z) = g_0(\mathbf{z}) \text{ qui est la fonction génératrice de } X(0). \quad (2.40)$$

1. On résout (2.39) pour $g_0(\mathbf{z})$ quelconque à partir des solutions particulières $G_i(t, \mathbf{z})$ avec la condition $G_i(0, \mathbf{z}) = z_i$ de la façon suivante :

$$g(t, z) = g_0(G_1(t, z), \dots, G_K(t, z))$$

car

$$\begin{aligned} \partial_t g(t, \mathbf{z}) &= \sum_j \partial_t g_0(t, \mathbf{z}) \partial_{z_j} g_0(\mathbf{G}(t, \mathbf{z})) \\ &= \sum_j \left(\sum_{l=1}^K V_l(\mathbf{z}) \partial_{z_l} G_j(t, \mathbf{z}) \partial_{z_j} g_0(\mathbf{G}(t, \mathbf{z})) \right) \\ &= \sum_{l=1}^K V_l(\mathbf{z}) \sum_{j=1}^K \partial_{z_l} G_j(t, \mathbf{z}) \partial_{z_j} g_0(\mathbf{G}(t, \mathbf{z})) \\ &= \sum_{l=1}^K V_l(\mathbf{z}) \partial_{z_l} g(t, \mathbf{z}). \end{aligned}$$

2. On veut donc trouver les solutions particulières G_1, \dots, G_K .

Pour cela, on suit la section 2.4.4. On pose : $\mathcal{I}(t, \mathbf{y}) = (\mathcal{I}_1, \dots, \mathcal{I}_K)(t, \mathbf{y})$, avec $\mathbf{y} = (y_1, \dots, y_K)$ des intégrales premières, indépendantes du système.

On considère l'application :

$$\begin{pmatrix} t \\ \mathbf{y} \end{pmatrix} \xrightarrow{\phi} \begin{pmatrix} t \\ \mathcal{I}^T(t, \mathbf{y}) \end{pmatrix},$$

avec

$$\partial_t \mathcal{I}^T(t, \mathbf{y}) = \begin{pmatrix} -V_1(\mathcal{I}_1 \cdots \mathcal{I}_K) \\ \vdots \\ -V_K(\mathcal{I}_1 \cdots \mathcal{I}_K) \end{pmatrix}.$$

La fonction $\mathcal{I}(t, \mathbf{y})$ est la solution de condition initiale $\mathcal{I}(0, \mathbf{y}) = \mathbf{y}$.

Sa matrice jacobienne est

$$J_\phi(t, \mathbf{y}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \partial_t \mathcal{I}_1(t, \mathbf{y}) & \partial_{y_1} \mathcal{I}_1(t, \mathbf{y}) & \cdots & \partial_{y_K} \mathcal{I}_1(t, \mathbf{y}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_t \mathcal{I}_K(t, \mathbf{y}) & \partial_{y_1} \mathcal{I}_K(t, \mathbf{y}) & \cdots & \partial_{y_K} \mathcal{I}_K(t, \mathbf{y}) \end{pmatrix}, \quad \text{matrice } (1+K) \times (1+K)$$

Sa réciproque est

$$\begin{pmatrix} t \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} t \\ G_1(t, \mathbf{z}) \\ \vdots \\ G_K(t, \mathbf{z}) \end{pmatrix} \xrightarrow{\psi} \begin{pmatrix} t \\ z_1 \\ \vdots \\ z_K \end{pmatrix},$$

dont la matrice Jacobienne

$$J_\psi(t, \mathbf{z}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \partial_t G_1(t, \mathbf{z}) & \partial_{z_1} G_1(t, \mathbf{z}) & \cdots & \partial_{z_K} G_1(t, \mathbf{z}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_t G_K(t, \mathbf{z}) & \partial_{z_1} G_K(t, \mathbf{z}) & \cdots & \partial_{z_K} G_K(t, \mathbf{z}) \end{pmatrix}, \quad \text{matrice } (1+K) \times (1+K),$$

vérifie $J_\psi(t, \mathbf{z}) = (J_\phi(t, \mathbf{y}))^{-1}$ avec $\begin{cases} \mathbf{z} = \mathcal{I}(t, \mathbf{y}) \\ \mathbf{y} = G(t, \mathbf{z}) \end{cases}$.

Par inversion d'une matrice en blocs de $J_\phi(t, \mathbf{y})$, on identifie

$$(J_\phi(t, \mathbf{y}))^{-1} = \begin{pmatrix} 1 & 0 \\ -(\partial_{\mathbf{y}} \mathcal{I})^{-1} \partial_t \mathcal{I} & (\partial_{\mathbf{y}} \mathcal{I})^{-1} \end{pmatrix}.$$

On a alors

$$\partial_t \mathbf{G}^T(t, \mathbf{z}) = -(\partial_{\mathbf{y}} \mathcal{I})^{-1} \partial_t \mathcal{I}^T \quad \text{et} \quad \partial_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) = J_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) = (\partial_{\mathbf{y}} \mathcal{I}(t, \mathbf{y}))^{-1}.$$

Cela signifie que

$$\partial_t \mathbf{G}^T(t, \mathbf{z}) = -J_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) \partial_t I^T(t, \mathbf{y}) = J_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) V(\mathcal{I})$$

autrement dit

$$\partial_t \mathbf{G}^T(t, \mathbf{z}) = J_{\mathbf{z}} \mathbf{G}(t, \mathbf{z}) V(\mathbf{z})$$

Pour t et z fixés, il existe \mathbf{y} , origine de l'unique trajectoire $t \mapsto \mathcal{I}(t, \mathbf{y})$ telle que $\mathcal{I}(t, \mathbf{y}) = \mathbf{z}$. Les coordonnées de l'origine $\mathbf{y} = (y_1, \dots, y_K)$ sont des fonctions $G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})$ qui vérifient $G_j(0, \mathbf{z}) = z_j$. Elles sont donc des intégrales premières et toute solution $g(t, z)$, qui est une fonction génératrice s'écrira sous la forme

$$g_0(G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})).$$

Chapitre 3

Some parametric families of single-type branching distributions

Sommaire

3.1	Introduction	2
3.2	Continuous Time Branching Processes	3
3.2.1	Some Notations and Reminders	3
3.2.2	Branching Processes	4
3.3	Fundamental Differential Equations Driving Branching Processes	6
3.3.1	Properties of V :	7
3.3.2	Solving the PDE	8
3.3.3	Recovering Transition Probabilities	9
3.3.4	Moment Equations	11
3.4	Examples of Parametric distributions for Branching Processes	13
3.4.1	Quadratic Branching	13
3.4.2	Cubic branching	16
3.4.3	Geometric Branching Distribution	17
3.4.4	Are Branching Processes Allowed to be Poisson Distributed?	19

Résumé

In this paper, we present some parametric families of probability distributions associated to particular single type homogeneous branching processes in continuous time. Their simplicity and the relevance of the interpretation of parameters for many domains of applications are of valuable interest for statistical inference. These families are particularly well adapted to handle branching dynamical systems of populations where Poisson assumption is generally but mistakenly assumed. Calculations and pertinent properties concerning these probability distributions are derived from their generating functions satisfying specific linear partial differential equations. Besides, these equations allow the statement of a general recurrence formula for factorial moments.

keyword : Continuous-time Branching Processes, Partial Differential Equations, probability generating function, factorial moments, zero inflated distributions.

3.1 Introduction

In its first version, the idea sustaining the definition of a branching process (BP) is the description of simple random mechanisms allowing an individual of a population to breed into several copies or to die. The population starts with an initial random number of individuals, then at any time and independently from the others, each living individual after an exponentially distributed lifetime gives rise to a progeny of random, possibly null, size. Additionally to independence, the random behaviors of individuals are identically distributed (i.i.d.). These processes are proven to be Markovian and therefore inherit many important properties. A standard reference on Markov branching processes is, among many others [23].

The theory of branching processes considers appropriate stochastic models to describe dynamical systems of population dynamics whose components (cells, particles, etc.) breed, evolve, interact or die. Indeed, branching processes have been successfully applied in diverse domains [37] highlighting therefore many issues in areas such that of molecular biology, cell biology, developmental biology, epidemiology, demography, particle physics, evolution, ecology, medicine, and others. However, the issue of statistical inference in all this variety of dynamics unveil a flagrant lack of the existence of consistent families of probability distributions modeling the time varying population size $X(t)$ of these very dynamics (see [58] for inferential items).

Much of the early work on the continuous-time branching processes was initiated by the Russian school in the middle of the XXth century[52]. Later, many profound theoretical results regarding the stochastic behavior of branching processes were established by an ever greater international community of probabilists. First, it was proven that, depending on the mean value of the offspring size m , discrete time and then continuous time branching processes has a dichotomous behavior in the sense that the population becomes extinct or explodes almost surely (a.s) :

$$\lim_{t \rightarrow \infty} X(t) = 0 \quad \text{or} \quad \lim_{t \rightarrow \infty} X(t) = +\infty.$$

BP are divided into subcritical ($m < 1$), critical ($m = 1$) and supercritical ($m > 1$) types. An important characteristic for BP is the extinction probability function $p_0(t) = \mathbb{P}(X(t) = 0)$, so that $\mathbb{P}(X(\infty) = 0) = p_0(\infty)$ and $\mathbb{P}(X(\infty) = \infty) = 1 - p_0(\infty)$. We prove in this paper that $p_0(t)$ actually determines the probability distribution of the whole process. On the other hand, general BP can explode at finite times, letting $\tau_\infty = \sup\{t > 0 : X_t < \infty\}$, ie the explosion time of the branching process, we call X regular if $\mathbb{P}(\tau_\infty < \infty) = 0$. Conditions for regularity are also well known (see [24], [23]) and were assumed in the present article.

This paper is mainly devoted to provide some sensible parametric distribution families of BP so that one can unroll the usual notwithstanding efficient maximum likelihood procedure in many applied areas. For that purpose, we followed an unusual approach based on the study of the partial differential equation (PDE) leading to the probability generating function (pgf) of BP. This approach revealed to be enough fruitful so that we obtained a general recurrence formula for factorial moments, when they exist. It also enlightens the bona fide relation between the extinction probability function and the pgf.

Chapter 2 introduces continuous time branching processes. Chapter 3 uses the Markovian property to derive the main differential and partial differential equations issued from Kolmogorov Equations. Chapter 4 applies the approach developed in Chapter 1 to investigate some particular branching dynamics.

3.2 Continuous Time Branching Processes

3.2.1 Some Notations and Reminders

First of all, let us define the subsequent notations used in this paper and remind some well known facts.

For enough smooth and well defined functions f, g , the n^{th} derivative of f is denoted $f^{(n)}(x)$ (or f', f'' if $n = 1, 2$), while the composition of f by g is written $f \circ g(x) = f(g(x))$ and $f^{\circ n}$ for the n -fold.

Recall that an integer valued r.v. Y with probability density (pdf) $\mathbb{P}(Y = j) = p_j, j \geq 0$, has a probability generating function (pgf) defined by

$$G(z) = \mathbb{E}(z^Y) = \sum_{j=0}^{\infty} p_j z^j$$

where $z \in D(0, 1)$, the complex open unit disc. G has good properties. For example G is analytic in D and as such is completely determined by its values on the real segment $[0, 1]$. It also inherits most of classical Fourier transforms properties. For example, Y has finite moments of order k if and only if $\lim_{z \uparrow 1} \partial_{z^k}^k G(1) = G^{(k)}(1^-) < \infty$. When this is the case, the factorial moments $m^{[k]} = \mathbb{E}(X(X-1) \cdots (X-k+1))$ are then given by

$$m^{[k]} = G^{(k)}(1^-), \quad k \geq 0.$$

For example $\mathbb{E}(Y) = G'(1)$, $Var(Y) = G''_Y(1) + G'_Y(1)(1 - G'_Y(1))$ if $k=2$. In duality, the derivatives of the pgf at $z = 0$, yield the formulas for probabilities

$$p_j = \frac{G^{(j)}(0)}{j!}.$$

3.2.2 Branching Processes

Continuous time Markov branching process $\mathbf{X} = \{X_t : t \geq 0\}$ are the analogous of discrete time Galton-Watson process $(X_n)_{n \geq 0}$, counting the number of individuals or particles of a population undergoing a branching dynamics. We briefly recall some of the definitions, tools and results related to the latter processes before dealing only with the continuous time context.

Discrete Time Branching Processes

Discrete time Markov branching processes $(X_n)_{n \geq 0}$ are defined by an initial integer valued r.v. X_0 , and by the following recursive rule. At time $n \geq 1$, if $X_n = 0$ then $X_{n+1} = 0$ a.s., whereas for $X_n > 0$, we have $X_{n+1} = \sum_{j=1}^{X_n} Y_{n,j}$, where the $Y_{n,j}$ are i.i.d. integer valued random variables with probability distribution $(\mathbb{P}(Y = j) = p_j, j \geq 0)$. This construction/definition implies that $(X_n)_{n \geq 0}$ is a homogeneous Markov process.

The random variable $Y_{n,j}$ represents the size of the offspring issued from the individual j living at time n . The event $Y_{n,j} = 0$ means the death of individual j at time n . The event $Y_{n,j} = 1$ can be interpreted without any loss of comprehension either as the survival or the *instantaneous restoration* of the individual j , for the sequel.

The complete independence and identity of the branching behaviors of particles entails that if $X(0)$ has pgf $g_0(\cdot)$ and if the $Y_{n,j}$ have the common pgf G , then the iterations of G yield the pgf of X_n :

$$g_n(z) = g_0 \circ G \circ \dots \circ G(z) = g_0 \circ G^{\circ n}(z).$$

The particular case $g_0(z) = z$ ie $X(0) = 1$ a.s, yields $g_n(z) = G^{\circ n}(z)$ and is actually sufficient to characterize all subsequent probability issues for the process.

Continuous Time Branching Processes

Homogeneous continuous time branching processes also assume that particles are independent, have the same offspring size distribution save that particles have continuous independent proper lifetimes that are i.i. exponentially distributed with parameter λ . The process $X(t)$, $t \geq 0$ counting the number of living particles at time t is also Markovian. Similarly to the discrete time case, the independence, homogeneity and Markov properties entail that the probability distribution of the process is entirely determined by the one variate distributions of the $X(t)$ conditionally to the event $X(0) = 1$ or equivalently by the probability generating function (pgf) $g(t, z) = \mathbb{E}(z^{X(t)} | X(0) = 1)$. For continuous time branching dynamics, the pgf reveals also to be the best tool for their theoretical investigation (see [3]).

For a consistent interpretation throughout the paper, we assume that at a death instant τ , the particle cannot be restored (reborn) instantly so that there is a genuine

jump for the population size. The other representation allowing instantaneous restoration of individuals is also possible but involves a distinct expression for the infinitesimal generator of the Markov process.

Being Markov and homogeneous, branching processes are entirely characterized by their transition probabilities, one may refer to [3, 23, 30] and for the general theory of continuous time jump Markov processes [24, 56] : $P(t) = (P_{i,j}(t) = \mathbb{P}(X(t) = j \mid X(0) = i), i, j \in \mathbb{N}), t \geq 0$. They satisfy the semi group property $P(t+s) = P(t)P(s)$. Under regularity conditions, P has an infinitesimal generator $Q = (Q_{i,j}, i, j \in \mathbb{N})$ (*i.e.* a derivative) defined by

$$Q_{i,j} = \lim_{t \rightarrow 0} \frac{P_{i,j}(t) - \delta_{\{i=j\}}}{t}.$$

For branching processes, the generator takes the following explicit form :

$$Q_{i,j} = \begin{cases} 0 & \text{if } j < i - 1 \\ i\lambda\mu & \text{if } j = i - 1 \\ -i\lambda & \text{if } j = i \\ i\lambda(1 - \mu)p_k & \text{if } j = i + k, \end{cases} \quad (3.1)$$

where $0 \leq \mu \leq 1$ and $\lambda \geq 0$. The evident degenerate cases $\mu = 0$, $\mu = 1$ and $\lambda = 0$ corresponding to almost surely monotonic or constant processes are elementary and are not considered thereafter.

The probability distribution of a regular branching process is therefore characterized by the pair (λ, π) where $\lambda > 0$ and $\pi = (q_j)_{0 \leq j \neq 1}$, a probability distribution on $\mathbb{N} - \{1\}$ such that

$$Q_{i,j} = i\lambda q_{j-i+1}, \quad j \neq i.$$

Kolmogorov's Equations

The infinite dimension matrix function of transition probabilities $(P(t), t \geq 0)$, is the unique solution of the backward Kolmogorov equation (BKE) :

$$\frac{dP_{i,j}(t)}{dt} = \sum_{k \geq 0} Q_{i,k} P_{k,j}(t) \quad (3.2)$$

Formally, equation (3.2) can be written as the linear infinite dimensional differential equation $P'(t) = QP(t)$ satisfying the initial condition $P(0) = I$. As for finite dimensional differential equations, its solution is written $P(t) = \exp(tQ) = \sum_{j \geq 0} t^j Q^j / j!$.

$P(t)$ commutes with Q , so that in the case of homogeneous Markov processes, $P(t)$ also obeys the linear forward Kolmogorov equation (FKE) : $\frac{dP_{i,j}(t)}{dt} = \sum_{k \geq 0} P_{i,k}(t) Q_{k,j}$, formally written $P'(t) = P(t)Q$.

Being infinite dimensional, both differential equations are difficult to handle in general. So instead of dealing with such transition probabilities, we resort to the more convenient generating function.

3.3 Fundamental Differential Equations Driving Branching Processes

Branching equations are based on the assumption that the sub-processes resulting from ancestors (*i.e.* branches) are independent and add to each other, therefore if $G(t, z)$ denote the pgf of $X(t)$ when $X(0) = 1$, we get $g(t, z) = g_0(G(t, z))$ if X_0 has pgf $g_0(z)$. Accordingly, one has only to investigate the case $G(t, z)$:

$$G(t, z) = \sum_{j=0}^{\infty} P_{1,j}(t)z^j, \quad z \in D(0, 1). \quad (3.3)$$

satisfying the initial condition $G(0, z) = z$.

The semi group property of $P(s+t) = P(t)P(s)$, can be translated into the property

$$G(t + s, z) = G(t, G(s, z)).$$

Taking account of the initial condition $G(0, z) = z$, we readily get

$$\begin{aligned} \partial_t G(t, z) &= \lim_{s \rightarrow 0} \frac{G(t + s, z) - G(t, z)}{s} \\ &= \lim_{s \rightarrow 0} \frac{G(t, G(s, z)) - G(t, z)}{s} \\ &= \lim_{s \rightarrow 0} \frac{G(s, z) - G(0, z)}{s} \frac{G(t, G(s, z)) - G(t, z)}{G(s, z) - z} \\ &= \partial_t G(0, z) \partial_z G(t, z), \end{aligned}$$

that is

$$\partial_t G(t, z) = \partial_t G(0, z) \partial_z G(t, z) \quad (3.4)$$

Next, from Equation 3.2 and 3.3, we deduce that

$$\partial_t G(t, z) = \sum_j Q_{1,j} G^j(t, z). \quad (3.5)$$

In particular, we get $\partial_t G(0, z) = \sum_j Q_{1,j} z^j$. For its importance we denote the latter expression

$$V(z) = \sum_{j \geq 0} Q_{1,j} z^j. \quad (3.6)$$

Eventually, Equation 3.4 will be written as the following fundamental partial derivative equation (PDE) for G :

$$\partial_t G(t, z) = V(z) \partial_z G(t, z). \quad (3.7)$$

3.3.1 Properties of V :

For interpretation purpose, V can be rewritten in a unique canonical form as follows :

$$V(z) = \lambda(G_Y(z) - z) \quad (3.8)$$

where $\lambda = -Q_{1,1}$ and $G_Y(z)$ is the pgf of the instantaneous offspring variable Y with values in $\mathbb{N} - \{1\}$ and

$$\mathbb{P}(Y = j) = -\frac{Q_{1,j}}{Q_{1,1}} = q_j, \quad 0 \leq j \neq 1. \quad (3.9)$$

It is convenient for application to think Y as the product $U(W + 2)$ of independent r.v. where U is Bernoulli with $\mathbb{P}(U = 0) = q_0$ and W has pdf

$$\mathbb{P}(W = j) = \frac{q_{j+2}}{1 - q_0}, \quad j \geq 0.$$

For instance W may be binomial, Poisson or geometric distributed.

The function V has the following properties on $[0, 1]$:

- V is analytic and satisfies $V(0) = \lambda q_0 > 0$ and $V(1) = 0$.
- $\mathbb{E}(Y) \geq 2(1 - q_0)$, since

$$\mathbb{E}(Y) = \sum_{j \geq 2} j q_j \geq 2 \sum_{j \geq 2} q_j = 2(1 - q_0).$$

- Its derivatives $V^{(k)}(z) = \lambda \sum_{m=0}^{\infty} \frac{(k+m)!}{m!} q_{k+m} z^m$, $k \geq 2$, satisfy in particular

$$V^{(k)}(0) = \lambda k! q_k, \quad V^{(k)}(1) = \lambda \sum_{m=0}^{\infty} \frac{(k+m)!}{m!} q_{k+m}. \quad (3.10)$$

- $V^{(k)}$ is non negative for $k \geq 2$, V and all of its derivatives are convex. V is strictly convex if $q_0 < 1$ and therefore has at most one zero z_V^* on the open real interval $]0, 1[$.

More precisely, we have the following factorization result :

Lemma 1. $V(z) = \lambda(1 - z) \left(q_0 - \sum_{j \geq 2} q_j (z + \dots + z^{j-1}) \right)$.

Moreover if $0 < q_0 < 1$, V has a (unique) zero within $]0, 1[$ if and only if $\mathbb{E}(Y) > 1$. This condition is satisfied in particular for $q_0 < 1/2$.

Proof.

Let us write z as $z = \sum_{j \neq 1} q_j z$ and factorize $z^j - z$ as usual to readily get

$$V(z) = \lambda(1 - z) \left(q_0 - \sum_{j \geq 2} q_j (z + \dots + z^{j-1}) \right).$$

Next, let us notice that $\theta(z) = \sum_{j \geq 2} q_j(z + \dots + z^{j-1})$ is strictly increasing on $[0, 1]$, satisfies $\theta(0) = 0$ and that

$$\theta(1) = \sum_{j \geq 2} q_j(j-1) = \sum_{j \geq 2} jq_j - \sum_{j \geq 2} q_j = \mathbb{E}(Y) - (1 - q_0).$$

Therefore, for V to have a zero within $[0, 1[$, it is necessary and sufficient that

$$q_0 - \theta(1) = 1 - \mathbb{E}(Y) < 0.$$

The last assertion results from the inequality $\mathbb{E}(Y) \geq 2(1 - q_0)$.

3.3.2 Solving the PDE

To solve the linear PDE 3.7, we have to exhibit a first integrals, that is, a function $G(t, z)$ that keeps constant on graphs $(t, \mathcal{I}(t, y))$, ie paths of solutions of the ordinary differential equation (ODE) :

$$\frac{d\mathcal{I}(t, y)}{dt} = -V(\mathcal{I}(t, y)), \quad \mathcal{I}(0, y) = y. \quad (3.11)$$

Indeed, if G is a first integral, then $v(t, y) \stackrel{\text{def}}{=} G(t, \mathcal{I}(t, y)) = G(0, y) = v(0, y)$ satisfies

$$\begin{aligned} \partial_t v(t, y) &= \partial_t G(t, \mathcal{I}(t, y)) + \partial_z G(t, \mathcal{I}(t, y)) \partial_t \mathcal{I}(t, y) \\ &= \partial_t G(t, \mathcal{I}(t, y)) - \partial_z G(t, \mathcal{I}(t, y)) V(\mathcal{I}(t, y)) = 0. \end{aligned}$$

Therefore, setting $\mathcal{I}(t, y) = z$ yields $\partial_t G(t, z) = V(z) \partial_z G(t, z)$ and this exactly corresponds to Equation (3.7).

Next, on the one hand, we notice that the more natural first integrals one may suggest consists in the determination of the origin coordinate y of solutions I passing through a point z at time t since this coordinates remains constant along the trajectory graph. This actually amounts to seek the inverse of the vector function :

$$\begin{pmatrix} t \\ y \end{pmatrix} \xrightarrow{F} \begin{pmatrix} t \\ \mathcal{I}(t, y) \end{pmatrix} = \begin{pmatrix} t \\ z \end{pmatrix} \xrightarrow{F^{-1}} \begin{pmatrix} t \\ G(t, z) \end{pmatrix} = \begin{pmatrix} t \\ y \end{pmatrix}$$

On the other hand, Equation (3.11), rewritten as

$$\frac{d\mathcal{I}}{V(\mathcal{I})} = -dt$$

has, on any interval not containing zeroes of V , primitives H that satisfy

$$H(\mathcal{I}) - H(\mathcal{I}_0) = \int_{\mathcal{I}_0}^{\mathcal{I}} \frac{du}{V(u)} = -t.$$

So, for any fixed t , if $\mathcal{I}(0, y) = y$, then $\mathcal{I}(t, y)$ is implicitly defined by :

$$H(\mathcal{I}(t, y)) = H(y) - t$$

Equivalently, the use of equalities $G(t, z) = y$ and $\mathcal{I}(t, y) = z$, proves that the pgf $G(t, z)$ satisfies the implicit equation :

$$H(G(t, z)) = H(z) + t. \quad (3.12)$$

Let us detail now the two cases related to the existence of a zero of V . Since V has at most one zero z_V^* in the interval $[0, 1[$, local inverses H^{-1} of H exist in both intervals $[0, z_V^*[$ and $]z_V^*, 1[$. Consequently, if these inverses are known by explicit expressions, we readily obtain

$$G(t, z) = H^{-1}(H(z) + t).$$

Otherwise, one can apply the implicit function theorem to get a Taylor series expansion at any order of G , or better to get, as we will see, the probability density of $X(t)$ by using the equality :

$$\partial_{z^k} H(G(t, z)) = \partial_{z^k} H(z). \quad (3.13)$$

Without loss of generality, one can assume that $H(0) = 0$, so that, setting $z = 0$ in 3.12, yields the following formula for the extinction probability $P_{1,0}(t) = G(t, 0)$:

$$H(P_{1,0}(t)) = t. \quad (3.14)$$

Corollary 1. *The primitive H of $1/V$ on the interval $[0, z_v^*[$ satisfying $H(0)=0$ is the increasing inverse function of the extinction probability function $P_{1,0}(t)$.*

Corollary 2. *$G(t, z)$ is entirely characterized by the extinction probability function, that is :*

$$G(t, z) = P_{1,0}(t + H(z)).$$

3.3.3 Recovering Transition Probabilities

In order to characterize the pdf $P_{1,n}(t) = \mathbb{P}(X(t) = n / X(0) = 1)$, $n \geq 0$ for any fixed t , we first recall that $P_{1,n}(t) = \frac{1}{n!} \partial_{z^n} G(t, 0)$ and then resort to Faà di Bruno's formula related to n -th derivatives of a function composition. Under the Bell polynomial version, the latter reads :

$$(f \circ g)^{(n)}(x) = \sum_{k=1}^n f^{(k)}(g(x)) \mathbf{B}_{n,k} \left(g^{(1)}(x), g^{(2)}(x), \dots, g^{(n-k+1)}(x) \right) \quad (3.15)$$

where

$$\mathbf{B}_{n,k}(u_1, \dots, u_{n-k+1}) = \sum_{j \in \sigma_{n,k}} \frac{n!}{j_1! \cdots j_{n-k+1}!} \left(\frac{u_1}{1!} \right)^{j_1} \cdots \left(\frac{u_{n-k+1}}{(n-k+1)!} \right)^{j_{n-k+1}}, \quad (3.16)$$

The sum over $j \in \sigma_{n,k}$, includes nonnegative integers $j = (j_1, j_2, \dots, j_{n-k+1})$ satisfying the constraints :

$$\begin{cases} j_1 + j_2 + \dots + j_{n-k+1} = k, \\ j_1 + 2j_2 + 3j_3 + \dots + (n-k+1)j_{n-k+1} = n. \end{cases} \quad (3.17)$$

We notice that when derivatives of the functions f and $f \circ g$ are known, Fa di Bruno Equation (3.15) generates a recursive relation for the derivatives of g .

Indeed, $g^{(n)}$ only appears for $k = 1$, and in that case, the sum in (3.16) reduces to a single index ($j_1 = 0, \dots, j_{n-1} = 0, j_n = 1$), that is $\mathbf{B}_{n,1}(u_1, u_2, \dots, u_n) = u_n$.

Consequently, on the set $\{x : f^{(1)}(g(x)) \neq 0\}$, we get the recurrence relation :

$$g^{(n)}(x) = \frac{(f \circ g)^{(n)}(x) - \sum_{k=2}^n f^{(k)}(g(x)) \mathbf{B}_{n,k}(g^{(1)}(x), \dots, g^{(n-k+1)}(x))}{f^{(1)}(g(x))} \quad (3.18)$$

In this latter context, let us fix t , put $f(z) = H(z)$ and $g(z) = G(t, z) = G_t(z)$ and notice that

$$(H \circ G_t)(z) = H(z) + t \Rightarrow (H \circ G_t)^{(n)}(z) = H^{(n)}(z).$$

So, the differentiation with respect to z yields

$$G_t^{(n)}(z) = \frac{H^{(n)}(z) - \sum_{k=2}^n H^{(k)}(G_t(z)) \mathbf{B}_{n,k}(G_t^{(1)}(z), \dots, G_t^{(n-k+1)}(z))}{H^{(1)}(G_t(z))}. \quad (3.19)$$

Consequently, if the inverse of H is not explicitly known, one can derive the Taylor expansion $G(t, z) = \sum_{n=0}^{\infty} G_t^{(n)}(0) z^n / n!$ via the recursive relation :

$$G_t^{(n)}(0) = V(G_t(0)) \left[H^{(n)}(0) - \sum_{k=2}^n H^{(k)}(G_t(0)) \mathbf{B}_{n,k}(G_t^{(1)}(0), \dots, G_t^{(n-k+1)}(0)) \right].$$

Moreover, $H^{(1)} = 1/V$ as a composition of V with the inverse function obeys the Fa di Bruno formula :

$$H^{(k)}(x) = \sum_{l=1}^{(k-1)} \frac{(-1)^l l!}{V(x)^{l+1}} B_{k-1,l}(V^{(1)}(x), \dots, V^{(k-l)}(x)), \quad k \geq 2.$$

This finally implies that the $P_{1,n}(t)$, $n \geq 0$ obey the recursive relation :

$$n! P_{1,n}(t) = V(P_{1,0}(t)) \left[H^{(n)}(0) - \sum_{k=2}^n H^{(k)}(P_{1,0}(t)) \mathbf{B}_{n,k}(1! P_{1,1}(t), \dots, (n-k+1)! P_{1,n-k+1}(t)) \right]. \quad (3.20)$$

The last formula shows how the extinction probability $P_{1,0}(t)$ completely determines the distribution probability of the branching process. The extinction probability itself, is completely determined by the coefficients $(Q_{1,j})_{j \geq 0}$ as follows.

Proposition 1. *The extinction probability $P_{1,0}(t)$ satisfies the equation $H(P_{1,0}(t)) = t$. The coefficients of its series expansion :*

$$P_{1,0}(t) = \sum_{n=1}^{\infty} \frac{P_{1,0}^{(n)}(0)}{n!} t^n,$$

satisfy the recurrence formula : $P_{1,0}^{(1)}(0) = Q_{1,0}$, and for $n \geq 2$

$$P_{1,0}^{(n)}(0) = -Q_{1,0} \sum_{n=2}^k H^{(k)}(0) \mathbf{B}_{n,k} \left(P_{1,0}^{(1)}(0), \dots, P_{1,0}^{(n-k+1)}(0) \right),$$

with

$$H^{(k)}(0) = \sum_{l=1}^{(k-1)} \frac{(-1)^l l!}{Q_{1,0}^{l+1}} B_{k-1,l} (1! Q_{1,1}, \dots, (k-l)! Q_{1,k-l}).$$

Proof.

Since $H(P_{1,0}(t)) = t$ and $P_{1,0}(0) = 0$, the differentiation with respect to t , yields

$$P_{1,0}^{(1)}(t) = 1/H^{(1)}(P_{1,0}(t)) = V(P_{1,0}(t)),$$

while for $n \geq 2$ we get $(H \circ P_{1,0})^{(n)}(t) = 0$. The recursive Fa di Bruno formula similar to equation (3.19), is written in this case :

$$H^{(1)}(P_{1,0}(t)) P_{1,0}^{(n)}(t) = - \sum_{k=2}^n H^{(k)}(P_{1,0}(t)) \mathbf{B}_{n,k} \left(P_{1,0}^{(1)}(t), \dots, P_{1,0}^{(n-k+1)}(t) \right).$$

Putting $t = 0$ and substituting 0 for x in $H^{(k)}(x)$ ends the proof. ■

Using Lemma (2), we get the following expression for the pgf :

Corollary 3.

$$G(t, z) = \sum_{n=1}^{\infty} \frac{P_{1,0}^{(n)}(0)}{n!} (H(z) + t)^n.$$

3.3.4 Moment Equations

The n^{th} -factorial moments of $X(t)$ defined by

$$m_{X(t)}^{[n]} = \mathbb{E} (X(t)(X(t) - 1) \cdots (X(t) - n + 1)), \quad n \geq 1,$$

always exist and, whether finite or infinite, satisfy the equality :

$$m_{X(t)}^{[n]} = \lim_{z \uparrow 1} \partial_{z^n} G(t, z) = \partial_{z^n} G(t, 1).$$

We have the following recursive formula for the factorial moments.

Proposition 2. *If $m_X^{[N]}(t) < \infty$, for some integer N , then, even if $\mathbb{E}(Y) - 1 = 0$, the following recursive integral formula holds true :*

$$m_X^{[1]}(t) = \exp^{\lambda(\mathbb{E}(Y)-1)t} \tag{3.21}$$

and, for $2 \leq n \leq N$,

$$m_X^{[n]}(t) = \lambda \sum_{k=2}^n \binom{n}{k} m_Y^{[k]} \int_0^t \exp^{n\lambda(\mathbb{E}(Y)-1)(t-s)} m_X^{[n-k+1]}(s) ds. \quad (3.22)$$

Proof. Using the Leibnitz rule for the derivatives of a product in Equation (3.7), yields (with some abuse of notation) :

$$\begin{aligned} \partial_t \partial_{z^n} G(t, z) &= \partial_{z^n} \partial_t G(t, z) = \partial_{z^n} (V(z) \partial_z G(t, z)) \\ &= \sum_{k=0}^n \binom{n}{k} V^{(k)}(z) G_t^{(n-k+1)}(z). \end{aligned}$$

So, if $n = 1$, the latter equation reduces to :

$$\partial_t \partial_z G(t, z) = V(z) \partial_{z^2} G(t, z) + V^{(1)}(z) \partial_z G(t, z).$$

Now, if we put $z = 1$ and observe that $V(1) = 0$, $V^{(1)}(1) = \lambda(\mathbb{E}(Y) - 1)$ and that $V^{(k)}(1) = \lambda m_Y^{[k]}$, where $m_Y^{[k]}$, $k > 1$ is the k^{th} -factorial moment of the instantaneous offspring variable Y . So, equating $m_X^{[k]}(t) = \partial_{z^k} G(t, 1)$ yields

$$\begin{aligned} \partial_t m_X^{[1]}(t) &= \lambda(\mathbb{E}(Y) - 1) m_X^{[1]}(t) \\ \partial_t m_X^{[n]}(t) &= \lambda n(\mathbb{E}(Y) - 1) m_X^{[n]}(t) + \lambda \sum_{k=2}^n \binom{n}{k} m_Y^{[k]} m_X^{[n-k+1]}(t) \end{aligned}$$

This is a recursive system of homogeneous (for $n = 1$) and non homogeneous (for $n > 1$) linear ODE for $m_X^{[n]}(t)$ when the $m_X^{[k]}(t)$, $1 \leq k \leq n - 1$ are previously determined. The corresponding solutions are classical and are respectively given by Equations (3.21) and (3.22) if we take account of the initial conditions $m_X^{[n]}(0) = \mathbf{1}_{\{n=1\}}$.

Note also that the previous formulas are still valid when $\mathbb{E}(Y) = 1$. ■

Since

$$\mathbb{E}(X(t)) = m^{[1]}(t) \text{ and } \text{Var}(X(t)) = m^{[2]}(t) + m^{[1]}(t) - \left(m^{[1]}(t)\right)^2,$$

we get in particular :

Corollary 4. For $\mathbb{E}(Y) - 1 \neq 0$,

$$\mathbb{E}(X(t)) = \exp^{\lambda(\mathbb{E}(Y)-1)t} \quad (3.23)$$

$$\text{Var}(X(t)) = \left(\frac{m_Y^{[2]}}{\mathbb{E}(Y) - 1} - 1 \right) \exp^{\lambda(\mathbb{E}(Y)-1)t} \left(\exp^{\lambda(\mathbb{E}(Y)-1)t} - 1 \right) \quad (3.24)$$

while, for $\mathbb{E}(Y) = 1$:

$$\mathbb{E}(X(t)) = 1, \quad \text{Var}(X(t)) = \lambda m_Y^{[2]} t. \quad (3.25)$$

3.4 Examples of Parametric distributions for Branching Processes

We now apply the results of the previous section to give a few examples of parametrized distributions of branching processes that can be of valuable interest in statistics issues.

3.4.1 Quadratic Branching

Apart the pure birth or death processes, the simplest homogeneous branching process requires only 2 parameters (λ, q) to define the infinitesimal transition rules :

$$Q_{1,j} = \begin{cases} \lambda q & \text{if } j = 0 \\ -\lambda & \text{if } j = 1 \\ \lambda(1 - q) & \text{if } j = 2 \end{cases}$$

So, we have in this case

$$V(z) = \lambda(q - z + (1 - q)z^2)$$

corresponds to a dichotomic instantaneous branching variable Y with pdf

$$\mathbb{P}(Y = 0) = q, \mathbb{P}(Y = 2) = 1 - q, \text{ and pgf } G_Y(z) = q + (1 - q)z^2.$$

This is why we call quadratic such a branching process. The factorial moments of Y are written

$$\mathbb{E}(Y) = m_Y^{[1]} = m_Y^{[2]} = 2(1 - q), \text{ whereas } m_Y^{[k]} = 0, k > 2.$$

The explosion condition $\mathbb{E}(Y) - 1 > 0$ corresponds here to the condition $q < 1/2$.

Searching adequate primitives of $1/V$

To integrate the rational function $1/V$, we have to distinguish the two cases depending of the single or multiple nature of the zeroes of V . Clearly, V has the factorization

$$V(z) = \lambda(1 - q)(1 - z) \left(\frac{q}{1 - q} - z \right).$$

The zeroes $z_1 = 1$ and $z_V = q/(1 - q)$ coincide iff $q = 1/2$.

Case A : $q = 1/2$ i.e. $z_V = 1$.

The primitive H of $1/V$, is defined for any $I \in [0, 1[$ by

$$H(\mathcal{I}) = \int_{\cdot}^{\mathcal{I}} \frac{dz}{V(z)} = \frac{2}{\lambda} \int_{\cdot}^{\mathcal{I}} \frac{dz}{(1 - z)^2} = \frac{2}{\lambda(1 - \mathcal{I})} + Cte.$$

Convening that $H(0) = 0$ ie $Cte = -\lambda/\mathcal{I}$, yields :

$$H(\mathcal{I}) = \frac{2\mathcal{I}}{\lambda(1-\mathcal{I})}.$$

H satisfies $\lim_{\mathcal{I} \uparrow 1} H(\mathcal{I}) = +\infty$ and has an inverse on $[0, 1[$:

$$H^{-1}(v) = \frac{\lambda v}{2 + \lambda v}, \quad v \geq 0.$$

Case B : $q \neq 1/2$.

The partial fraction expansion of $1/V$ is given by

$$\frac{1}{V} = \frac{1}{\lambda(2q-1)} \left(\frac{1}{1-z} - \frac{1}{z_V - z} \right).$$

Any primitive of $1/V$, defined on any adequate sub interval of $[0, 1[$ is written :

$$H(\mathcal{I}) = \int^{\mathcal{I}} \frac{dz}{V(z)} = \frac{1}{\lambda(2q-1)} \log \left(\frac{|z_V - \mathcal{I}|}{|1 - \mathcal{I}|} \right) + Cte.$$

Regarding the location of the root z_V with respect to the interval $[0, 1[$, we distinguish two cases :

Case B1. The extinction case $z_V > 1$ i.e. $q > 1/2$.

H is well defined and is an increasing function on $[0, 1[$. Convening that $H(0) = 0$, we obtain

$$H(\mathcal{I}) = \frac{1}{\lambda(2q-1)} \left[\log \left(\frac{z_V - \mathcal{I}}{1 - \mathcal{I}} \right) - \log z_V \right]$$

Again, $\lim_{\mathcal{I} \uparrow 1} H(\mathcal{I}) = +\infty$, and on $[0, 1[$ H has the inverse :

$$H^{-1}(v) = \frac{1 - \exp^{\lambda(2q-1)v}}{\frac{1}{z_V} - \exp^{\lambda(2q-1)v}}, \quad v \geq 0.$$

One has to notice that the exponential coefficient $\lambda(2q-1)$ is positive in this case.

Case B2. The explosive case $z_V < 1$ ie $q < 1/2$.

- On $D_1 = [0, z_V[$, H increases and verifies $\lim_{\mathcal{I} \uparrow z_V} H(\mathcal{I}) = +\infty$. Again, convening that $H(0) = 0$ we obtain the same formulas for H and H^{-1} .

- On $D_2 =]z_V, 1[$, H is decreases and satisfies $\lim_{\mathcal{I} \downarrow z_V} H(\mathcal{I}) = +\infty$ whereas

$$\lim_{\mathcal{I} \uparrow 1} H(\mathcal{I}) = -\infty.$$

Moreover, V attains it minimum at the inflexion point $z^* = 1/2(1-q) \in [z_V, 1[$ which satisfies $\log \frac{z_V - z^*}{1 - z^*} = 0$. So we can choose for the primitive satisfying $H(z^*) = 0$, that is :

$$H(\mathcal{I}) = \frac{1}{\lambda(2q-1)} \log \left(\frac{\mathcal{I} - z_V}{1 - \mathcal{I}} \right)$$

whose inverse is :

$$H^{-1}(v) = \frac{z_V + \exp^{\lambda(2q-1)v}}{1 + \exp^{\lambda(2q-1)v}}, \quad v \in \mathbb{R}.$$

Pgf and pdf of $X(t)$

Before stating the exact form of the pgf and the pdf of $X(t)$, we give the following definition.

Definition 1. An integer valued random variable N is said to be a zero inflated geometric distribution with parameters (α, β) , $\alpha \geq 0$, $\beta \leq 1$, if it is the mixture of a geometric distribution and a Dirac mass at 0, that is :

$$\begin{aligned}\mathbb{P}(N = 0) &= 1 - \alpha \\ \mathbb{P}(N = k) &= \alpha\beta(1 - \beta)^{k-1}, \quad k \geq 1\end{aligned}$$

A zero inflated geometric distribution has the following characteristics :

$$\begin{aligned}G_N(z) &= \frac{(1 - \alpha) - ((1 - \beta - \alpha)z)}{1 - (1 - \beta)z}, \\ \mathbb{E}(N) &= \frac{\alpha}{\beta} \text{ and } Var(N) = \frac{\alpha(2 - \beta - \alpha)}{\beta^2}.\end{aligned}$$

Proposition 3. $X(t)$ has a zero inflated geometric distribution with parameters (α_t, β_t) , ie

$$G(t, z) = \frac{(1 - \alpha_t) - (1 - \beta_t - \alpha_t)z}{1 - (1 - \beta_t)z}$$

where :

1. $\alpha_t = \beta_t = \frac{2}{\lambda t + 2}$, if $q = 1/2$,
2. and for $q \neq 1$ and $z_V = q/(1 - q)$:

$$\alpha_t = \frac{1 - z_V}{1 - z_V \exp^{\lambda(2q-1)t}}, \quad \beta_t = \frac{(1 - z_V) \exp^{\lambda(2(q-1)t)}}{1 - z_V \exp^{\lambda(2(q-1)t)}}.$$

Proof. One has only to apply the formula $G(t, z) = H^{-1}(H(z) + t)$ in the different cases (A, B1, and B2) and notice that whatever is the sub-case, all formulas yield the same expression. ■

Corollary 5. In particular, a quadratic branching process $X(t)$ with parameters (λ, q) satisfies :

1. $\mathbb{P}(X(t) = 0) = 1 - \alpha_t$
2. $\mathbb{E}(X(t)) = 1$, $Var(X(t)) = \lambda t$ for $q = 1/2$,
3. and for $q \neq 1$:

$$\mathbb{E}(X(t)) = \exp^{-\lambda(2q-1)t}, \quad Var(X(t)) = \frac{\exp^{-\lambda(2q-1)t}(\exp^{-\lambda(2q-1)t} - 1)}{1 - 2q}.$$

3.4.2 Cubic branching

The cubic branching process depending on 3 parameters, is defined by $V(z) = \lambda(G_Y(z) - z)$, via the Y -pgf :

$$G_Y(z) = q_0 + q_2 z^2 + q_3 z^3,$$

under the constraints : $q_0 + q_2 + q_3 = 1$, and $0 < q_0, q_3 < 1$.

The offspring distribution satisfies

$$m_Y^{[1]} = 2q_2 + 3q_3, \quad m_Y^{[2]} = 2q_2 + 6q_3, \quad m_Y^{[3]} = 6q_3, \quad \text{and } m_Y^{[k]} = 0, k \geq 4$$

So we have in particular

$$\mathbb{E}(Y) = 2q_2 + 3q_3 \quad \text{and} \quad \text{Var}(Y) = (3q_0 + q_2)(3q_3 + 2q_2) - 2q_2.$$

The explosion condition $\mathbb{E}(Y) - 1 > 0$ ie $q_2 + 2q_3 > q_0$, is also equivalent to say that a zero of V belongs to $]0, 1[$.

Indeed, the polynomial factorization of V gives

$$V(z) = \lambda(z - 1) \left(q_3 z^2 + (1 - q_0)z - q_0 \right) = \lambda q_3 (z - 1)(z - z_+)(z - z_-),$$

with

$$z_{\pm} = \frac{(q_0 - 1) \pm \sqrt{(1 - q_0)^2 + 4q_0 q_3}}{2q_3}.$$

Clearly, the parameter constraints imply that $z_- < 0, z_+ > 0$, then using Lemma 1, we see that V has a zero $z_V = z_+$ in $]0, 1[$ (resp. $z_+ = 1$) iff $q_2 + 2q_3 > q_0$ (resp. $q_2 + 2q_3 = q_0$).

Moment Formulas

Applying Proposition 2 readily yields the following recursive formula.

Proposition 4. *According to the extinction/explosion case, the two first moments ($k = 1, 2$) are written :*

1. for $2q_2 + 3q_3 = 1$, $\mathbb{E}(X(t)) = 1$, $\text{Var}(X(t)) = \lambda(1 + 3q_3)t$,
2. for $2q_2 + 3q_3 \neq 1$: $\mathbb{E}(X(t)) = \exp^{\lambda(2q_2 + 3q_3 - 1)t}$,

$$\text{Var}(X(t)) = \frac{3q_3 + 1}{2q_2 + 3q_3 - 1} \exp^{\lambda(2q_2 + 3q_3 - 1)t} \left(\exp^{\lambda(2q_2 + 3q_3 - 1)t} - 1 \right).$$

Whereas for higher moments $k \geq 3$, we get

$$m_X^{[n]}(t) = \lambda n(n-1) \int_0^t \exp^{n\lambda(2q_2 + 3q_3 - 1)(t-s)} \left((q_2 + 3q_3)m_X^{[n-1]}(s) + (n-2)q_3 m_X^{[n-2]}(s) \right) ds.$$

Probability Generating functions

We now consider the form of the primitive H regarding the location of the zero z_+ of V .

Case A : $z_+ \neq 1$. The three zeroes are distinct, so taking $H(0) = 0$ yields

$$H(\mathcal{I}) = \int_{\cdot}^{\mathcal{I}} \frac{dz}{V(z)} = \frac{1}{\lambda q_3} \log \left(|\mathcal{I} - 1|^A |\mathcal{I} - z_-|^B |\mathcal{I} - z_+|^C \right),$$

where

$$A = -(B + C), \quad B = \frac{1}{(1 - z_-)(z_+ - z_-)}, \quad C = \frac{1}{(z_+ - 1)(z_+ - z_-)},$$

that is,

$$H(\mathcal{I}) = \frac{1}{\lambda q_3 (z_+ - z_-)} \left(\frac{1}{(1 - z_-)} \log \left(\left| 1 + \frac{1 - z_-}{\mathcal{I} - 1} \right| \right) - \frac{1}{(1 - z_+)} \log \left(\left| 1 + \frac{1 - z_+}{\mathcal{I} - 1} \right| \right) \right).$$

Case B : $z_+ = 1$ corresponds to a double root of V , and then

$$\frac{1}{V(z)} = \frac{1}{\lambda q_3} \left(\frac{A}{(z - z_-)} + \frac{B}{(z - 1)} + \frac{C}{(z - 1)^2} \right),$$

where

$$A = \frac{1}{(1 - z_-)^2}, \quad B = -\frac{1}{(1 - z_-)^2}, \quad C = \frac{1}{(1 - z_-)}.$$

Therefore, taking $H(0) = 0$, yields

$$H(\mathcal{I}) = \frac{1}{\lambda q_3} \left(\log(|\mathcal{I} - z_-|^A |\mathcal{I} - 1|^B) - C(\mathcal{I} - 1)^{-1} \right).$$

Moreover in the case B, we have $1 - z_- = (3 + q_3^{-1})/2$ which ultimately yields :

$$H(\mathcal{I}) = \frac{1}{\lambda q_3 (1 - z_-)^2} \left(\log \left(\left| 1 + \frac{1 - z_-}{\mathcal{I} - 1} \right| \right) + \frac{1 - z_-}{\mathcal{I} - 1} \right)$$

Remark.

For both cases, save for some specific parameter values, the inverse of H cannot be expressed by classical functions. However H is simple enough to allow fast computational procedures for inversion (see example in Figure ??). In the same vein, H is also enough simple to fully develop the recursive relations given in Proposition 1. So that one can approximate the extinction probability function $P_{1,0}(t)$ at any order and therefore get the pgf via the formula $G(t, z) = P_{1,0}(t + H(z))$.

3.4.3 Geometric Branching Distribution

This is a 3-parametrized (λ, q_0, α) family of branching distributions where the instantaneous branching distribution takes the form

$$G_Y(z) = q_0 + (1 - q_0)z^2 \frac{\alpha}{1 - (1 - \alpha)z}.$$

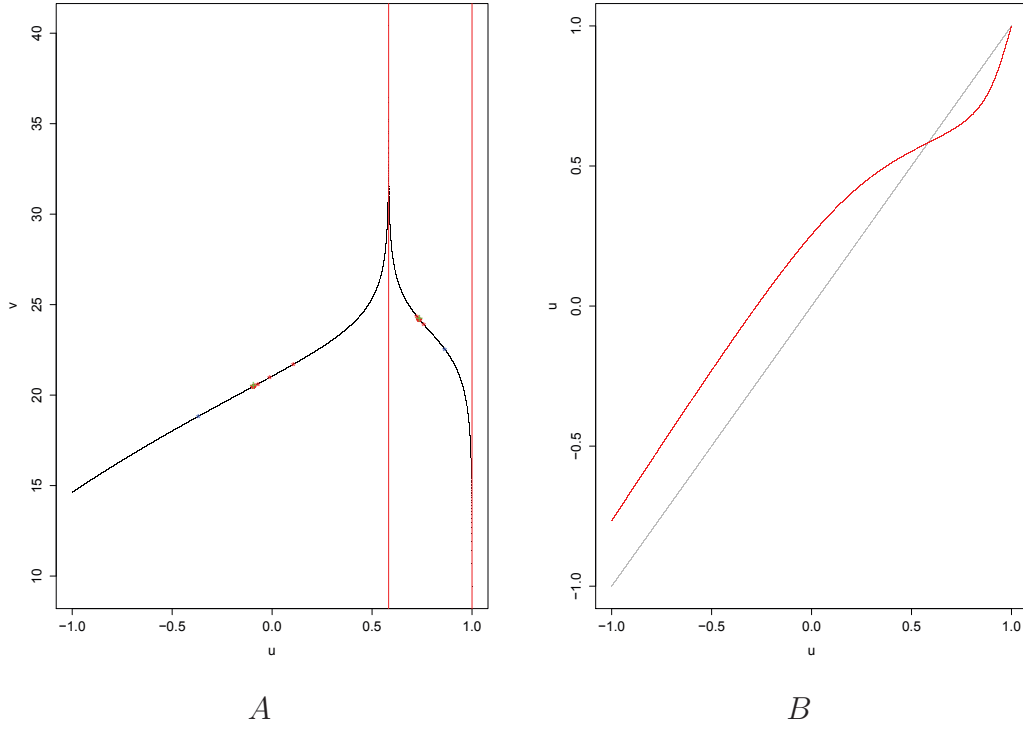


FIGURE 3.1 – **A.** Graph of the primitive $H(z)$ for a cubic branching process. **B.** A numeric approximation of the function H^{-1} allows the computation of the pgf $G(5, z) = H^{-1}(H(z) + 5)$, $z \in [-1, 1]$, for a cubic branching at a time $t = 5$.

As already pointed out Y can be written $Y = U(2+W)$ with independent U and W . U is Bernoulli distributed $B(1 - q_0)$ and W is geometrically distributed with probabilities

$$\mathbb{P}(W = k) = (1 - \alpha)^k \alpha, \quad k \geq 0.$$

For this family, we have $m_Y^{[1]} = (1 - q_0)(1 + \alpha)/\alpha$, and for $k \geq 2$

$$m_Y^{[k]} = \alpha(1 - q_0)(1 - \alpha)^{(k-2)} \sum_{l \geq 0} \frac{(k+l)!}{l!} (1 - \alpha)^l.$$

and in particular, $Var(Y) = (1 - q_0)(4 + (1 - \alpha)(2 + \alpha)/\alpha^2)$.

The function $V(z) = \lambda(G_Y(z) - z)$ expresses as a rational function :

$$V(z) = \lambda \frac{q_0 - (1 + q_0(1 - \alpha))z + (1 - q_0\alpha)z^2}{1 - (1 - \alpha)z},$$

that can be factorized into

$$V(z) = \tilde{\lambda} \frac{(z - 1)(z - z_v)}{z - z_*},$$

where $\tilde{\lambda} = \frac{\lambda(1 - q_0\alpha)}{\alpha - 1}$, $z_v = \frac{q_0}{1 - \alpha q_0}$ and $z_* = \frac{1}{1 - \alpha}$.

In addition to $z_* > 1$, we notice that $\alpha > 0$, and $q_0 \neq 1$ imply that $z_* \neq 1$, $z_* \neq z_v$, so that we have to deal only with the following two cases.

Case $z_V = 1$ corresponds to equality $q_0 = 1/(1 + \alpha)$ so that 1 is a double root. Choosing a primitive H with $H(0) = 0$ yields

$$H(\mathcal{I}) = \frac{1}{\tilde{\lambda}} \int_0^{\mathcal{I}} \frac{dz}{V(z)} = \frac{1}{\tilde{\lambda}} \left(\log |\mathcal{I} - 1| - \frac{1 - z_*}{\mathcal{I} - 1} \right)$$

H is well defined and increases on $[0, 1[$. It equals the convex combination

$$H(\mathcal{I}) = -\frac{1}{\lambda(1 - q_0\alpha)} \left((1 - \alpha) \log(1 - \mathcal{I}) + \alpha \frac{1}{1 - \mathcal{I}} \right)$$

and has an inverse on $[0, +\infty[$, precisely the extinction probability $P_{1,0}(t)$.

Case $z_V \neq 1$. We have not to distinguish between the extinction and non extinction cases $z_V > 1$ and $z_V < 1$, since for both we obtain the formula :

$$H(\mathcal{I}) = \frac{1}{\tilde{\lambda}} \int_0^{\mathcal{I}} \frac{dz}{V(z)} = \frac{1}{\tilde{\lambda}} \left[\left(1 - \frac{z_* - z_V}{1 - z_V} \right) \log |\mathcal{I} - 1| + \frac{z_* - z_V}{1 - z_V} \log |\mathcal{I} - z_V| \right].$$

In particular, on the interval $[0, 1 \wedge z_V[$, H is written

$$H(\mathcal{I}) = \frac{1}{\tilde{\lambda}} \log \left((z_V - \mathcal{I})^\theta (1 - \mathcal{I})^{1-\theta} \right)$$

with $\theta = \frac{z_* - z_V}{1 - z_V}$ and $\tilde{\lambda} = \lambda q_0 \frac{z_*}{z_V}$.

For both, the inverse of H is not a classical function, but H is enough simple to allow numeric approximations of H^{-1} or better than that to use the different recursive formulas for transition probabilities.

3.4.4 Are Branching Processes Allowed to be Poisson Distributed ?

In this example, we will not deal with the interesting example of Poisson distributions for the offspring. We are simply interested here to argue against the frequent use of Poisson or zero inflated Poisson distributions in applications related to branching processes. Indeed, the Poisson distributions, possibly zero inflated, occur only for very specific simple branching dynamics. First, let us give the corresponding definition.

Definition 2. An integer valued r.v. X is said to be zero inflated Poisson distributed i.e. $X \sim ZIP(\alpha, \nu)$, $0 \leq \alpha \leq 1$ and $\nu > 0$, if

$$\begin{aligned} \mathbb{P}(X = 0) &= 1 - \alpha \\ \mathbb{P}(X = n) &= \frac{\alpha}{1 - e^{-\nu}} \frac{e^{-\nu} \nu^n}{n!}, \quad n \geq 1. \end{aligned}$$

The generating function of X has the form :

$$G_X(z) = 1 - \alpha \frac{e^{\nu(z-1)} - 1}{e^{-\nu} - 1}.$$

Proposition 5. *A regular branching process $X(t)$ with infinitesimal generator $(Q_{i,j}(\lambda, \mu, p_k, k \geq 1))$ given by Equation (3.1), has marginal distributions that are zero inflated Poisson $ZIP(\alpha(t), \nu(t))$ for some non negative smooth functions $0 \leq \alpha(t) \leq 1$ and $\nu(t) \geq 0$ if and only if it corresponds to the degenerate case of a pure death process (ie $\mu = 1$). In that case, we have $X(0) \sim ZIP(\alpha(0), \nu(0))$, $\nu(t) = \nu(0)e^{-\lambda t}$ and $\alpha(t) = \alpha(0) \frac{e^{-\nu(t)} - 1}{e^{-\nu(0)} - 1}$.*

Proof For sake of simplicity, let us denote $\alpha_t = \alpha(t)$. Next, let us assume that the marginal distribution of $X(t)$ is $ZIP(\alpha_t, \nu_t)$ distributed, with pgf

$$g(t, z) = 1 - \alpha_t \frac{e^{\nu_t(z-1)} - 1}{e^{-\nu_t} - 1}. \quad (3.26)$$

First, the smoothness assumptions imply that the initial condition is also of the same type, that is $g(0, \mathbf{z}) = 1 - \alpha_0 \frac{e^{\nu_0(z-1)} - 1}{e^{-\nu_0} - 1}$.

Next, since $X(t)$ is a branching process, it satisfies the fundamental equation, $\partial_t g(t, \mathbf{z}) = V(z) \partial_z g(t, z)$, which is thereby written :

$$\frac{\alpha'_t}{\alpha_t} (e^{\nu_t(z-1)} - 1) + \nu'_t \left((z-1)e^{\nu_t(z-1)} + e^{-\nu_t} \frac{e^{\nu_t(z-1)} - 1}{e^{-\nu_t} - 1} \right) = V(z) \nu_t e^{\nu_t(z-1)}.$$

Setting $\theta_t = \frac{\alpha'_t}{\alpha_t} e^{\nu_t} + \frac{\nu'_t}{e^{-\nu_t} - 1}$ yields

$$V(z) = \frac{1}{\nu_t} \left(\frac{\alpha'_t}{\alpha_t} - \frac{\nu'_t}{e^{-\nu_t} - 1} - \theta_t e^{-\nu_t z} + \nu'_t z \right).$$

The series expansion of $e^{-\nu_t z}$ added to the canonical form $V(z) = \lambda(G_Y(z) - z)$ in Equations 3.8 and 3.9, where $G_Y(z) = q_0 + \sum_{k=2}^{\infty} q_k z^k$, assert that

$$\begin{aligned} \lambda &= -\frac{\nu'_t + \theta_t \nu_t}{\nu_t} \\ q_0 &= \frac{\alpha'_t e^{\nu_t} - 1}{\alpha_t \nu'_t + \theta_t \nu_t} \\ q_k &= \frac{\theta_t}{\nu'_t + \theta_t \nu_t} \frac{(-\nu_t)^k}{k!}, \quad k \geq 2 \end{aligned}$$

are all non negative and independent from t . However, this is possible only for $\theta_t \equiv 0$, otherwise the q_k would have opposite signs for even and odd k .

As a consequence, we obtain $\frac{\alpha'_t}{\alpha_t} = \frac{-\nu'_t e^{-\nu_t}}{e^{-\nu_t} - 1}$, which, taking account of the initial conditions, is written : $\alpha_t = \alpha_0 \frac{e^{-\nu_t} - 1}{e^{-\nu_0} - 1}$.

In this latter case, one readily gets $q_0 = 1$, meaning that the death parameter μ in the infinitesimal generator equals 1, corresponding therefore to a pure death process and that necessarily, we have $\nu_t = \nu_0 e^{-\lambda t}$.

Note that $\lambda = -\nu'_t / \nu_t = \nu'_0 / \nu_0$ translates the time scale in play within the branching dynamics. ■

Chapitre 4

Une famille paramétrique de lois de branchement multi-types

Sommaire

4.1	Introduction	22
4.2	Simulation de quelques comportements de branchements multi-types	23
4.2.1	Populations indépendantes	23
4.2.2	Populations en interaction	24
4.3	Description du modèle du branchement multi-types	27
4.4	Une famille particulière paramétrique de lois de branchement multi-types	28
4.4.1	De l'EDP linéaire au système d'EDO autonomes associé au branchement	29
4.4.2	Réduction de la solution du système d'EDO de la 2 ^{ème} partie (4.6)	30
4.4.3	Résolution de la première partie de (4.7)	31
4.4.4	Résolution numérique des EDO	39
4.5	Un exemple de branchement 3-types	41
4.5.1	Simulations de l'exemple de branchement 1 mère-2 filles	46
4.6	Sous-modèle du branchement multi-types mère-filles	48
4.7	Un exemple du sous-modèle : branchement 4-types	52
4.7.1	Simulations de l'exemple du sous-modèle : branchement 4-types	57
4.8	Branchement uni-type issu d'un branchement multi-types	60

4.1 Introduction

Les processus de branchement multi-types sont des modèles stochastiques appropriés pour décrire les systèmes dynamiques de populations en interaction, ainsi que certaines dynamiques en physique des particules. Ils jouissent d'un large éventail d'applications qui recouvrent des domaines aussi variés que la biologie (voir [45, 64, 40]), l'épidémiologie, la démographie mais aussi la chimie et la physique des particules.

Autrement dit, on décrit l'évolution de K populations (ou classes) d'individus en interaction dite de branchement. Cela signifie que chaque individu évolue au cours du temps indépendamment de tous les autres et de sa date de naissance (*i.e.* son âge) et qu'au bout d'un laps de temps aléatoire (de loi propre à sa classe), il meurt, ou se reproduit en plusieurs individus répartis dans les différentes classes ; un individu d'une classe k vit un temps exponentiel de paramètre λ_k et donne naissance à plusieurs individus, qui se rajoutent aux autres, selon une loi multidimensionnelle $p_k(dn_1, \dots, dn_k)$. L'indépendance et l'équidistribution au sein d'une même classe est une hypothèse fondamentale pour le développement d'outils mathématiques adéquats à l'étude de ces processus de branchements.

Dans ce chapitre, nous présentons des familles paramétriques de lois associées à des processus de branchement multi-types, en temps continu et homogènes en temps. Par la recherche de la simplicité et la pertinence du paramétrage, ces familles s'avèrent adaptées à la description de systèmes dynamiques plus maniables pour les applications. Les calculs et propriétés concernant leur lois de probabilité utilisent comme en unidimensionnel les fonctions génératrices correspondantes et se rapportent à la résolution de certaines équations linéaires aux dérivées partielles.

On rappelle qu'un processus de branchement multi-types classique est un processus de Markov homogène $\mathbf{X}(t) = (X_1(t), \dots, X_K(t)) \in \mathbb{N}^K$, où $\mathbf{X}(t)$ décrit le nombre d'individus de chaque type à l'instant t .

Rappelons aussi quelques notations utilisées tout au long de ce chapitre. Ainsi, on considère pour une variable complexe $\mathbf{z} = (z_1, \dots, z_K) \in D(0, 1)^K$, où $D(0, 1)^K$ est le disque ouvert complexe dans \mathbb{C}^K et pour un multi-indice $i = (i_1, \dots, i_K) \in \mathbb{N}^K$, on écrira $\mathbf{z}^i = z_1^{i_1} \times \dots \times z_K^{i_K}$, ainsi que $i! = i_1! \times i_2! \times \dots \times i_K!$ et $\partial_{\mathbf{z}}^i = \partial_{z_1^{i_1} \times \dots \times z_K^{i_K}}$. La transposée d'une matrice ou d'un vecteur V sera notée V^T . Pour étudier la loi de ces processus, ayant une condition initiale aléatoire, il est suffisant d'utiliser l'indépendance des individus.

Il est pratique de considérer dans un premier temps les fonctions génératrices de la loi du processus $\mathbf{X}(t)$, lorsque le processus démarre avec un seul individu de type k (voir [2, 22, 51]), on notera :

$$G_k(t, \mathbf{z}) = \mathbb{E} \left(\mathbf{z}^{\mathbf{X}(t)} \mid \mathbf{X}(0) = \mathbf{e}_k \right), \quad k = 1, \dots, K$$

et $\mathbf{G}(t, \mathbf{z}) = (G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z}))$.

Ainsi, si $\mathbf{X}(0)$ a une loi $p_0 = (dn_1, \dots, dn_K)$ de génération $g_0(\mathbf{z})$ alors

$$g(t, \mathbf{z}) = \mathbb{E}(z^{\mathbf{X}(t)}) = g_0(G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z}))$$

On rappelle aussi que les dérivées partielles d'une fonction génératrice $g(\mathbf{z})$ de variable Y à valeurs dans \mathbb{N}^K , déterminent entièrement sa loi de probabilité par la formule :

$$\mathbb{P}(Y = i) = \frac{1}{i!} \partial_{\mathbf{z}^i} g(\mathbf{0})$$

De même, on sait que Y a un moment d'ordre $r = (r_1, \dots, r_K)$, c'est à dire que $\mathbb{E}(Y_1^{i_1}, \dots, Y_K^{i_K}) < \infty$, si et seulement si $\lim_{\mathbf{z} \nearrow \mathbf{1}} \partial_r g(\mathbf{z})$ est finie.

4.2 Simulation de quelques comportements de branchements multi-types

4.2.1 Populations indépendantes

Exemple 1 : Cas explosif

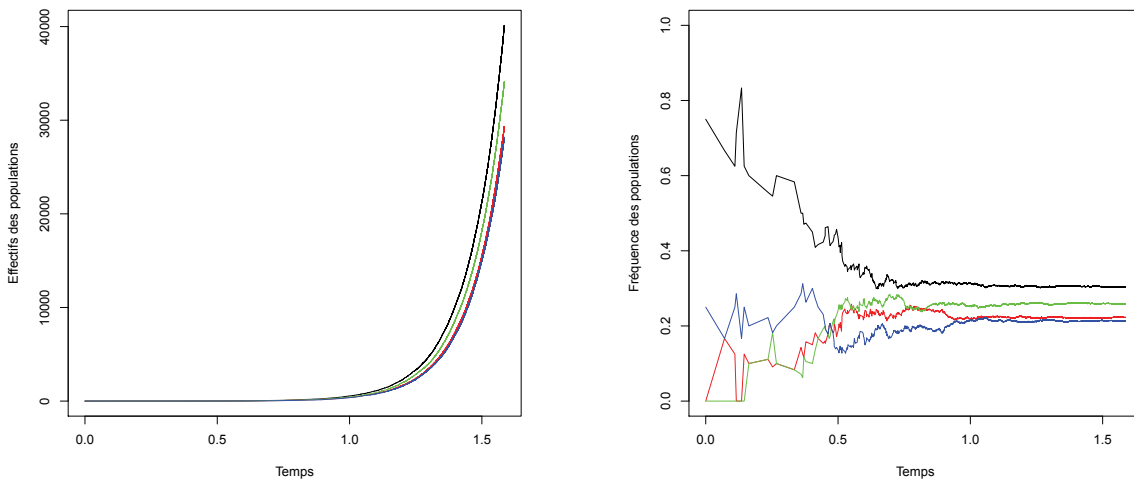


FIGURE 4.1 – Cas explosif de 4 populations indépendantes.

Simulations de branchement avec 4 populations indépendantes. Les graphes de la figure 4.1 représentent les effectifs (à gauche) et les fréquences relatives (à droite) de chaque population au cours du temps, lors d'une simulation à 10 000 branchements. La taille initiale de chaque population est aléatoire selon la loi de Poisson de paramètre 1. Le paramètre de temps d'attente λ exponentiel de base avant branchement pour chaque population suit une loi uniforme sur $[0.001, 1]$. Dans ce modèle, un individu d'une population i meurt ou se transforme en deux individus. Il s'agit en fait ici de copies indépendantes de branchements uni-types (voir la section 2.3.1 du chapitre 2).

Exemple 1 : Cas mixte

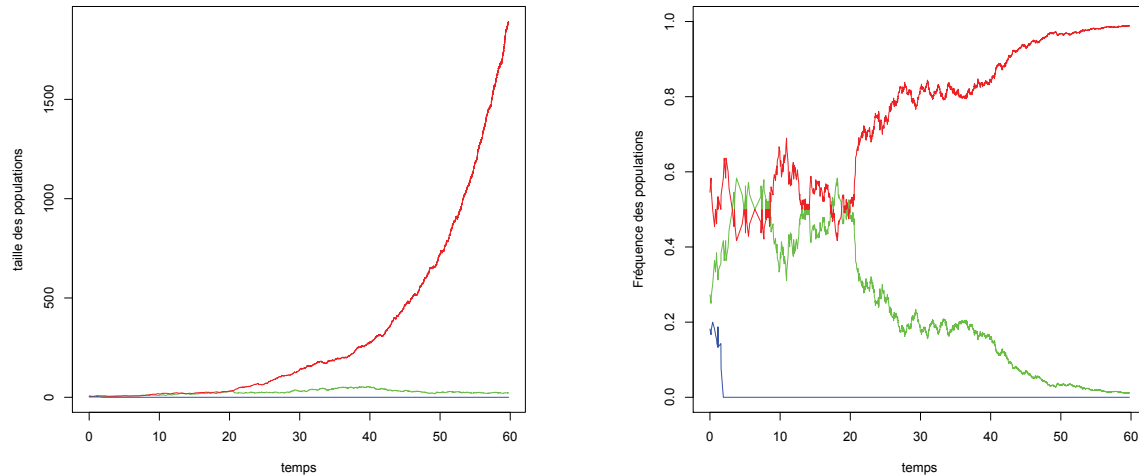


FIGURE 4.2 – Simulation des tailles dynamiques de 3 populations indépendantes en comportement mixte : explosion | extinction.

Le graphe 4.2 représente les effectifs (à gauche) et les fréquences (à droite) de 3 populations indépendantes les unes des autres au cours du temps. Ici la probabilité de mort μ varie pour chaque population (la population en rouge $\mu = 0.4$, en vert $\mu = 0.5$ et en bleu $\mu = 0.7$). On remarquera que le cas d’explosif (rouge) survient avec $\mu > \frac{1}{2}$, et le cas d’extinction pour les autres populations (bleu et vert) avec $\mu > \frac{1}{2}$ (voir la section 2.3.1 du chapitre 2).

4.2.2 Populations en interaction

Exemple 1 : Cas équitable (interaction en tous sens)

La figure 4.3 montre l’évolution de 3 populations, interagissant entre elles en tous sens (simulation de 10000 branchements). Le graphe représente les effectifs (à gauche) et les fréquences relatives (à droite) de chaque population au cours du temps. La probabilité de mort d’un individu, lors d’un saut est commune aux populations, et est notée μ . Dans ce modèle les populations interagissent en tous sens : un individu d’une population i meurt ou se transforme en deux individus de n’importe quelle population. On remarque que les 3 populations explosent mais gardent des proportions “stationnaires”.

Pour les graphes (a, b), on prend des lois de Poisson de paramètre égal à 3 pour la taille initiale de chaque population. Ici, $X_1(0) = 3$ (vert), $X_2(0) = 2$ (rouge), $X_3(0) = 3$ (bleu).

Le paramètre de temps d’attente avant branchement pour chaque population suit une loi uniforme sur $[0.001, 1]$. Ici, $\lambda_1 = 0.3732$, $\lambda_2 = 0.106$, $\lambda_3 = 0.657$. On remarque que

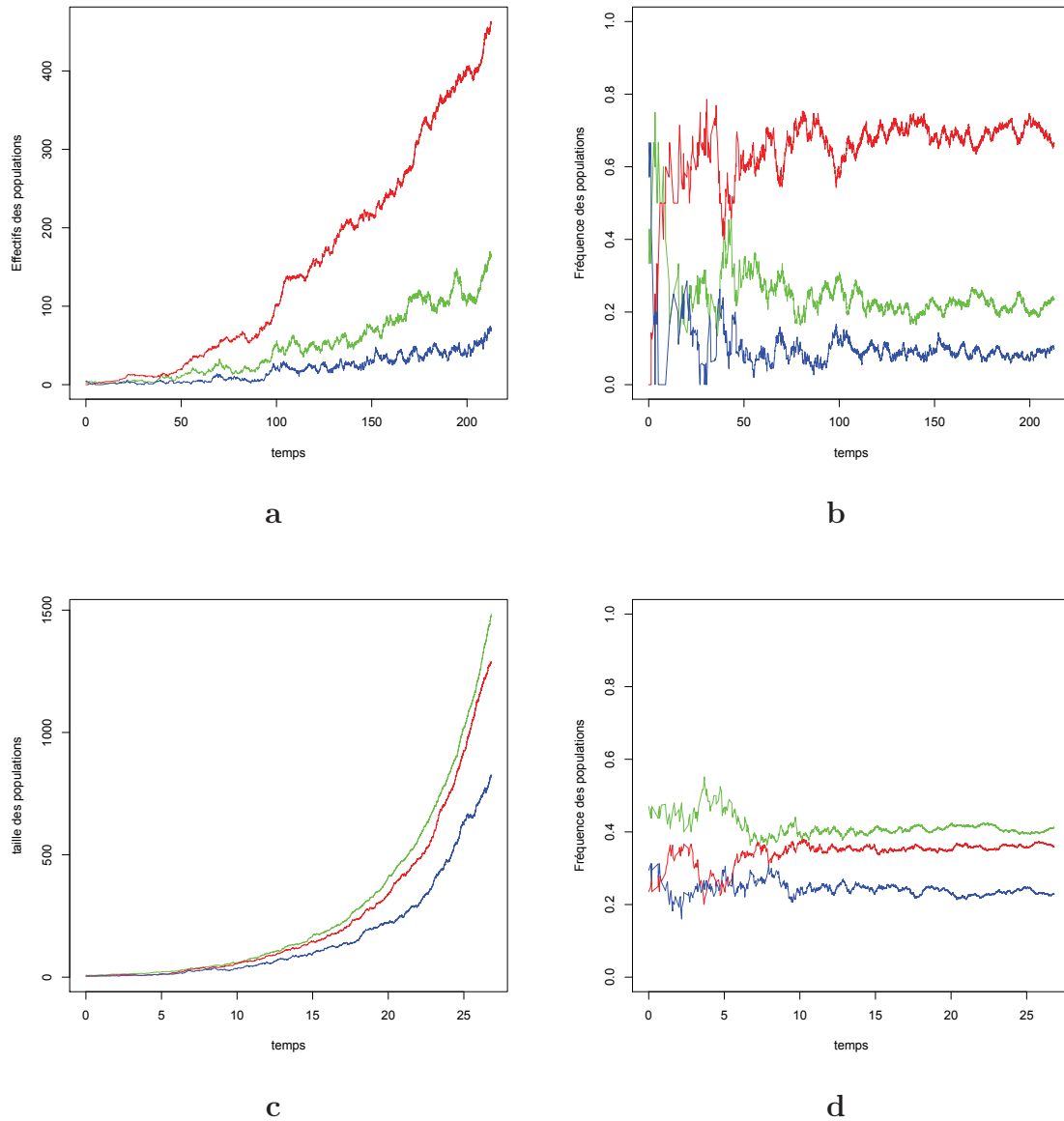


FIGURE 4.3 – Simulation des tailles de 3 populations en interactions équitables.

les moyennes de temps avant branchement sont $1/\lambda_3 < 1/\lambda_1 < 1/\lambda_2$ et on explique les ratios dans (b).

La probabilité de mort d'un individu lors d'un saut est $\mu = 0.6$, valeur commune aux 3 populations. On remarquera aussi que, pour $\mu > 1/2$, cela n'entraîne pas l'extinction automatique des populations.

Pour les graphes (c, d), le paramètre des lois de Poisson vaut 6 pour la taille initiale de chaque population. Dans cette simulation $X_1(0) = 8$ (vert), $X_2(0) = 7$ (rouge), $X_3(0) = 4$ (bleu) et le paramètre de temps d'attente suit une loi uniforme sur $[0, 1]$ pour chaque population. Ici, $\lambda_1 = 0.334$, $\lambda_2 = 0.497$, $\lambda_3 = 0.99$ et donc $1/\lambda_3 < 1/\lambda_2 < 1/\lambda_1$ justifiant les ratios observés dans (d).

Ici, la probabilité de mort d'un individu lors d'un saut commun aux populations vaut

$\mu = 0.5$ et cela n'entraîne pas non plus l'extinction.

Exemple 2 : Cas d'interactions hiérarchiques (à sens unique)

La Figure 4.4 représente le comportement de 3 populations qui interagissent selon un schéma hiérarchique. Dans ce modèle, la 1^{ère} population (vert) contribue aux 3 populations, la 2^{ème} (rouge) contribue à elle même et à la 3^{ème} (bleu) et la dernière (bleu) ne donne qu'à elle même.

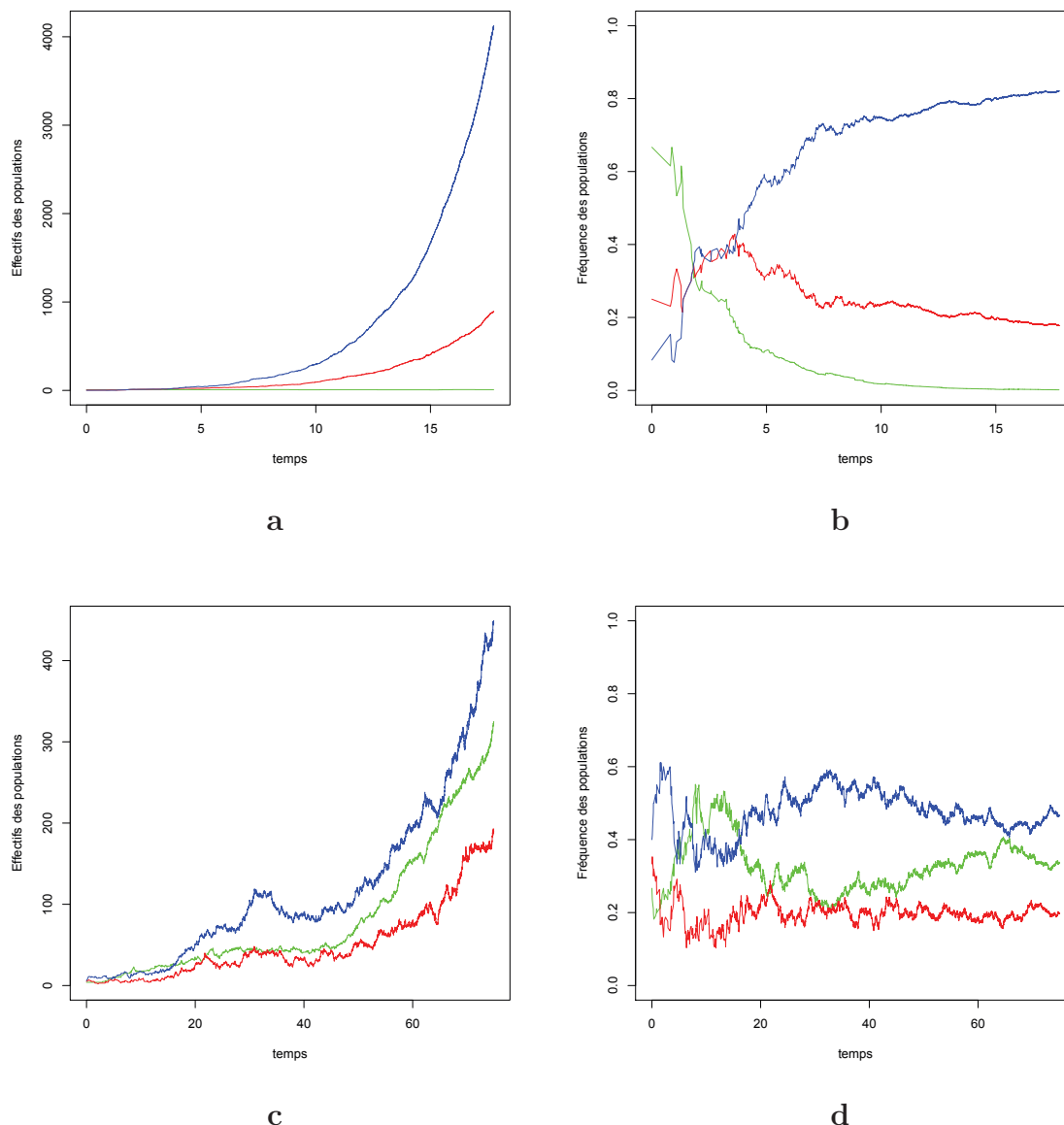


FIGURE 4.4 – Simulation des tailles de 3 populations en interactions hiérarchique.

Pour les graphes (a, b), la taille initiale de chaque population est aléatoire et de loi de Poisson de paramètre 4. Dans cette simulation, $X_1(0) = 4$, $X_2(0) = 7$, $X_3(0) = 3$.

Le paramètre de temps d'attente suit la loi uniforme sur $[10^{-3}, 1]$. Ici, $\lambda_1 = 0.979$, $\lambda_2 =$

0.825, $\lambda_3 = 0.285$, et le paramètre de décès $\mu = 0.3$ est commun à tous les individus. On remarquera que la 3^{ème} population receveuse universelle (bleu) évolue plus vite que la 2^{ème} population (rouge) mais dans un rapport assez stable. Par contre, la 1^{ère} population (vert) finit par mourir en dépit de la valeur du paramètre $\mu < 0.5$.

Pour les graphes (\mathbf{c}, \mathbf{d}) , les tailles initiales des populations valent $X_1(0) = 5$, $X_2(0) = 1$, $X_3(0) = 4$. Les paramètres des temps d'attente sont $\lambda_1 = 0.272$, $\lambda_2 = 0.439$ et $\lambda_3 = 0.548$. On remarquera que c'est le cas explosif pour les 3 populations. La 1^{ère} population (vert) donneuse a la plus petite probabilité de mort d'un individu lors d'un saut ($\mu_1 = 0.3$). Dans ce cas, elle a moins de risque d'extinction. Par contre les deux autres populations ont une probabilité de mort d'un individu lors d'un saut plus élevée ($\mu_2 = 0.6$, $\mu_3 = 0.7$), qui est finalement compensée par le rajout d'individus essentiellement issus de la classe 1.

4.3 Description du modèle du branchement multi-types

L'équation de Chapman-Kolmogorov (backward) pour les probabilités de transition $(P_{i,j}(t))$ d'un processus de branchement multi-types et homogène dans le temps s'écrit :

$$\partial_t P_{i,1}(t) = \sum_{l \in \mathbb{N}^K} Q_{i,1} P_{1,j}(t)$$

via le générateur infinitésimal $\mathbf{Q} = (Q_{i,l})$. Cette équation traduit fondamentalement la propriété de semi-groupe $\mathbf{G}(t + s, \mathbf{z}) = \mathbf{G}(t, \mathbf{G}(s, \mathbf{z}))$ avec $s, t \geq 0$ et $\mathbf{z} \in D(0, 1)^K$.

Ainsi pour tout $k = 1, \dots, K$, on a d'une part la définition et l'égalité suivante (voir chapitre 2)

$$V_k(\mathbf{z}) = \partial_t G_k(0, \mathbf{z}) = \sum_{\mathbf{l} \in \mathbb{N}^K} Q_{e_k, \mathbf{l}} \mathbf{z}^{\mathbf{l}}, \quad (4.1)$$

et d'autre part $G_k(t, \mathbf{z})$ qui vérifie l'EDP avec condition initiale :

$$\begin{cases} \partial_t G_k(t, \mathbf{z}) &= \sum_{h=1}^K V_h(\mathbf{z}) \partial_{z_h} G_k(t, \mathbf{z}) \\ G_k(0, \mathbf{z}) &= z_k. \end{cases} \quad (4.2)$$

On s'aperçoit que ces K-EDP à (K+1)-variables se réduisent à la résolution d'une seule EDP pour une fonction scalaire $g(t, z)$ mais avec des conditions initiales distinctes.

La solution usuelle du système linéaire ((4.2)) (voir chapitre 2) consiste à exhiber K intégrales premières indépendantes G_1, \dots, G_K pour le système ((4.2)), et d'en déduire ainsi que toute solution est de la forme

$$g(t, \mathbf{z}) = g_0(G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})),$$

où g_0 est déterminée par la condition initiale $g(0, \mathbf{z}) = g_0(\mathbf{z})$.

Pour cela, posons $\mathbf{W}^T(t, \mathbf{z}) = (1, -\mathbf{V}^T(\mathbf{z}))$ avec $\mathbf{V}^T(\mathbf{z}) = (V_1(\mathbf{z}), \dots, V_K(\mathbf{z}))$ et $\nabla = (\partial_t, \partial_{z_1}, \dots, \partial_{z_K})$ et remarquons que l'équation ((4.2)) s'écrit

$$\partial_{\mathbf{W}}G(t, \mathbf{z}) = \langle \mathbf{W}, \nabla G(t, \mathbf{z}) \rangle \equiv 0, \quad (4.3)$$

signifiant par là que $G(t, \mathbf{z})$ est constante le long des trajectoires $t \rightarrow \mathcal{I}(t, \mathbf{z})$ du système différentiel autonome K -dimensionnel

$$\frac{d\mathcal{I}}{dt} = -\mathbf{V}(\mathcal{I}(t)), \quad \mathcal{I}(0, \mathbf{y}) = \mathbf{y} \quad (4.4)$$

Formellement, il est facile de déterminer des intégrales premières indépendantes en remarquant que par le théorème d'unicité de la solution des EDO, les coordonnées à l'origine d'une trajectoire solution sont indépendantes et ne changent pas le long d'une trajectoire. Ce sont donc des intégrales premières !

Il s'agit donc des fonctions $G_i(t, \mathbf{z}) = y_i$ telles que $\mathcal{I}_i(t, \mathbf{y}) = \mathbf{z}$, avec la condition initiale $\mathcal{I}(0, \mathbf{y}) = \mathbf{y}$. Les $G_k(0, \mathbf{z}) = z_k$, $k = 1, \dots, K$, sont donc les K intégrales premières indépendantes.

De cette façon, on prouve que toute solution scalaire $g(t, \mathbf{z})$ de ((4.2)) avec la condition initiale $g(0, \mathbf{z}) = g_0(\mathbf{z})$ aura la forme $g(t, \mathbf{z}) = g_0(G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z}))$.

4.4 Une famille particulière paramétrique de lois de branchement multi-types

Nous considérerons un cas particulier à la fois "traitable", mais aussi suffisamment riche en comportements pour être utile pour des applications étudiant le développement de populations en interaction. On particularise l'équation générale (4.2) en rajoutant une population spéciale avec un indice $j = 0$, c.à.d.

$$\partial_t \mathbf{G}(t, \hat{\mathbf{z}}) - \sum_{j=0}^K V_j(\hat{\mathbf{z}}) \partial_{\hat{\mathbf{z}}_j} \mathbf{G}(t, \hat{\mathbf{z}}) = 0, \quad \hat{\mathbf{z}} = (z_0, z_1, \dots, z_K) = (z_0, \mathbf{z})$$

Ainsi on aura $K + 1$ populations où la première population notée de type "0" jouera le rôle spécial de population des *mères*, comme par exemple des cellules souches qui se reproduisent/meurent de façon indifférenciée dans un premier temps. Cette classe "donneuse" évoluera de façon autonome par rapport aux autres avec une dynamique simple de reproduction quadratique pour elle même et en fournissant linéairement les K autres classes, c.à.d. :

$$V_0(\hat{\mathbf{z}}) = \mu_0 - \Lambda_0 z_0 + \lambda_{00} z_0^2 + \lambda_{01} z_1 + \dots + \lambda_{0K} z_K, \quad \Lambda_0 = \mu_0 + \sum_{j=0}^K \lambda_{0j}.$$

Les K autres classes interagissent aussi, mais de façon linéaire entre elles, autrement dit,

$$\begin{cases} V_1(\mathbf{z}) = \mu_1 - \Lambda_1 z_1 + \lambda_{12} z_2 + \cdots + \lambda_{1K} z_K \\ \vdots \\ V_K(\mathbf{z}) = \mu_K + \lambda_{K1} z_1 + \cdots + \lambda_{K,K-1} z_{K-1} - \Lambda_K z_K \end{cases}, \quad \Lambda_i = \mu_i + \sum_{j \neq i}^K \lambda_{ij}.$$

En remarquant que V_1, \dots, V_K ne dépendent pas de z_0 , le système devient :

$$\partial_t G_i(t, \hat{\mathbf{z}}) = \sum_{j=0}^K V_j(\hat{\mathbf{z}}) \partial_{\hat{\mathbf{z}}_j} G_i(t, \hat{\mathbf{z}}), \quad i = 0, \dots, K.$$

4.4.1 De l'EDP linéaire au système d'EDO autonomes associé au branchement

D'après la section 2.4.5 du chapitre 2, en prenant $\hat{\mathbf{y}} = (y_0, \mathbf{y})$, $\mathbf{y} = (y_1, \dots, y_K)$, il faudra résoudre le système d'EDO suivant :

$$\frac{d}{dt} \begin{pmatrix} \mathcal{I}_0(t, \hat{\mathbf{y}}) \\ \mathcal{I}_1(t, \mathbf{y}) \\ \vdots \\ \mathcal{I}_K(t, \mathbf{y}) \end{pmatrix} = \begin{pmatrix} \Lambda_0 & -\lambda_{01} & \cdots & -\lambda_{0K} \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \mathcal{I}_0(t, \hat{\mathbf{y}}) \\ \mathcal{I}_1(t, \mathbf{y}) \\ \vdots \\ \mathcal{I}_K(t, \mathbf{y}) \end{pmatrix} - \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_K \end{pmatrix} + \begin{pmatrix} -\lambda_{00} \mathcal{I}_0^2(t, \hat{\mathbf{y}}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (4.5)$$

avec A la matrice

$$A = \begin{pmatrix} \Lambda_1 & -\lambda_{12} & \cdots & -\lambda_{1K} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ -\lambda_{K1} & \cdots & -\lambda_{K,K-1} & \Lambda_K \end{pmatrix}.$$

a) Résolution de la deuxième partie du système (4.5)

On commence par résoudre la deuxième partie de l'équation, indépendante de $\mathcal{I}_0(t, \hat{\mathbf{y}})$ c'est à dire en considérant :

$$\mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1(t, \mathbf{y}), \dots, \mathcal{I}_K(t, \mathbf{y})), \quad \text{avec } \mathbf{y} = (y_1, \dots, y_K) \quad \text{indépendante de } y_0$$

et avec la condition initiale $\mathcal{I}(0, \mathbf{y}) = \mathbf{y}$.

Sa solution générale est classique. Elle est de la forme :

$$\mathcal{I}(t, \mathbf{y}) = e^{At} \left(\mathbf{y} + \int_0^t e^{-As} ds \nu_0 \right) \quad (4.6)$$

avec $\nu_0^T = -(\mu_1, \dots, \mu_K)$.

b) **EDO associée à la 1^{ème} partie du système (4.5)**

Si on pose

$$H(t, \mathbf{y}) = \mu_0 + \sum_{j=1}^K \lambda_{0j} \mathcal{I}_j(t, \mathbf{y}) = \mu_0 + (\lambda_{01}, \dots, \lambda_{0K}) e^{At} \left(\mathbf{y} + \left(\int_0^t e^{-As} ds \right) \nu_0 \right),$$

alors, la première partie unidimensionnelle de l'équation devient :

$$\begin{cases} \frac{d\mathcal{I}_0}{dt}(t, \hat{\mathbf{y}}) = \Lambda_0 \mathcal{I}_0(t, \hat{\mathbf{y}}) - \lambda_{00} \mathcal{I}_0^2(t, \hat{\mathbf{y}}) - H(t, \mathbf{y}) \\ \mathcal{I}_0(0, \hat{\mathbf{y}}) = y_0 \end{cases} \quad (4.7)$$

avec la contrainte sur les paramètres : $\mu_0 + \sum_{j=1}^K \lambda_{0j} = \Lambda_0 - \lambda_{00}$

4.4.2 Réduction de la solution du système d'EDO de la 2^{ème} partie (4.6)

Revenons à la solution de (4.6). Elle passe par la décomposition de A en blocs de Jordan (disons r blocs) qui permet de révéler les différents types d'interactions et la possibilité de les reparamétriser utilement.

Ainsi, il existe une matrice P inversible telle que $J = PAP^{-1}$ soit diagonale par blocs, comme ci dessous.

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}.$$

Cela permet d'écrire $e^{-As} = P^{-1}e^{-Js}P$. Il suffira alors de résoudre séparément le problème pour chaque bloc de Jordan. Ensuite, en prenant $\tilde{\mathcal{I}}(t, \tilde{\mathbf{y}}) = P\mathcal{I}(t, \mathbf{y})$, on voit qu'il vérifie $\frac{d\tilde{\mathcal{I}}(t, \tilde{\mathbf{y}})}{dt} = J\tilde{\mathcal{I}}(t, \tilde{\mathbf{y}}) + \tilde{\nu}_0$, avec $\tilde{\nu}_0 = P\nu_0$ et $\tilde{\mathbf{y}} = P\mathbf{y}$. Ainsi $\tilde{\mathcal{I}}(t, \tilde{\mathbf{y}})$ se décomposera en r composantes indépendantes. Rappelons qu'à chaque valeur propre α de A peut être associée à un ou plusieurs blocs de Jordan. Chaque bloc J_L , pour $L = 1, \dots, r$, de taille m_L , s'écrit sous la forme suivante

$$J_L = \begin{pmatrix} \alpha_L & 1 & 0 & \dots & 0 \\ 0 & \alpha_L & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \alpha_L \end{pmatrix} = \alpha_L \mathbb{I}_{m_L} + \mathbf{E}_L.$$

Rappelons aussi que \mathbf{E}_L est nilpotente, car $\mathbf{E}_L^h = 0$, si $h \geq m_L$. Ainsi,

$$e^{-J_L s} = e^{-\alpha_L s} \sum_{l=0}^{m_L-1} \frac{(-s)^l \mathbf{E}_L^l}{l!} = e^{-\alpha_L s} \begin{pmatrix} 1 & -s & \frac{(-s)^2}{2!} & \cdots & \frac{(-s)^{m_L-1}}{(m_L-1)!} \\ 0 & 1 & -s & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \frac{(-s)^2}{2!} \\ & & & \ddots & -s \\ 0 & \cdots & 0 & & 1 \end{pmatrix}, \quad (4.8)$$

et donc au final, pour un bloc J_L , la solution intégrale vectorielle sera :

$$\tilde{\mathcal{I}}_L(t, \mathbf{y}) = e^{J_L t} \left(\tilde{\mathcal{I}}_L(0, \mathbf{y}) + \int_0^t e^{-J_L s} ds \tilde{\nu}_{0L} \right), \quad (4.9)$$

avec $\tilde{\mathcal{I}}_L(0) = \tilde{\mathbf{y}}_L$, et $\tilde{\nu}_{0L}$ étant les composantes associées au bloc L . En regroupant les blocs, c.à.d. en utilisant

$$e^{Jt} = \begin{pmatrix} e^{J_1 t} & & \\ & \ddots & \\ & & e^{J_r t} \end{pmatrix},$$

la solution (4.6) devient :

$$\mathcal{I}(t, \mathbf{y}) = P^{-1} e^{Jt} \left(\tilde{\mathbf{y}} + \int_0^t e^{-J s} ds \tilde{\nu}_0 \right), \quad (4.10)$$

avec $\tilde{\mathbf{y}} = P\mathbf{y}$ et $\tilde{\nu}_0 = P\nu_0$.

4.4.3 Résolution de la première partie de (4.7)

Il reste enfin à trouver la solution de l'EDO (4.7), avec

$$H(t, \mathbf{y}) = \mu_0 + \tilde{\lambda}^T e^{Jt} \left(\tilde{\mathbf{y}} + \left(\int_0^t e^{-J s} ds \right) \tilde{\nu}_0 \right),$$

avec $\tilde{\lambda}^T = (\lambda_{01}, \dots, \lambda_{0K})P^{-1}$.

- i) On peut développer l'expression de $e^{Jt} \int_0^t e^{-J s} ds$. En omettant provisoirement l'indice de bloc L , la matrice (4.8), re-écrite $(\theta_{ij}(-s))_{i,j=1\dots m_l}$, donnera $e^{At} = (\theta_{ij}(t))$ et $e^{-At} = (\theta_{ij}(-t))$ avec :

$$\theta_{ij}(t) = \begin{cases} 0 & \text{si } i > j, \\ e^{\alpha t} \frac{t^{j-1}}{(j-1)!} & \text{si } i \leq j. \end{cases}$$

En notant $\int_0^t e^{-J s} ds = (\tilde{\theta}_{ij}(t))$, on a :

$$\tilde{\theta}_{ij}(t) = \int_0^t \theta_{ij}(-s) ds = \begin{cases} 0 & \text{si } i > j, \\ \frac{(-1)^{j-i}}{(j-i)!} \int_0^t e^{-\alpha s} s^{j-i} ds & \text{si } i \leq j. \end{cases}$$

Si $i \leq j$, on aura :

$$\begin{aligned}\tilde{\theta}_{ij}(t) &= \int_0^t \theta_{ij}(-s) ds \\ &= \frac{(-1)^{j-i}}{(j-i)!} \frac{e^{-\alpha t}}{(-\alpha)^{j-i+1}} [(-\alpha t)^{j-i} - (j-i)(-\alpha t)^{j-i-1} + (j-i)(j-i-1)(-\alpha t)^{j-i-2} \\ &\quad + \dots + (-1)^{j-i-1}(j-i)!(-\alpha t) + (-1)^{j-i}(j-i)!] - \frac{(-1)^{j-i+1}}{(\alpha)^{j-i+1}} \\ &= \frac{(-1)^{j-1}}{\alpha^{j-i+1}} \left[e^{-\alpha t} \sum_{l=0}^{j-i} \frac{(\alpha t)^l}{l!} - 1 \right].\end{aligned}$$

Pour tous $i = 1, \dots, m$, et $m \geq j \geq i$, on a :

$$\tilde{\theta}_{ij}(t) = \frac{(-1)^{j-1}}{\alpha^{j-i+1}} \left[e^{-\alpha t} \sum_{l=0}^{j-i} \frac{(\alpha t)^l}{l!} - 1 \right],$$

d'où

$$\tilde{\theta}_{ij}(t) = \begin{cases} (-\alpha)^{-1}[e^{-\alpha t} - 1] & \text{si } j = i, \\ (-\alpha)^{-2}[e^{-\alpha t}(1 + \alpha t) - 1] & \text{si } j = i + 1, \\ \vdots \\ (-\alpha)^{m_L-i-1}[e^{-\alpha t}(1 + \alpha t + \frac{(\alpha t)^2}{2!} + \dots + \frac{(\alpha t)^{m_L-i}}{(m_L-i)!}) - 1] & \text{si } j = m_L. \end{cases}$$

On peut aussi l'écrire sous la forme :

$$\tilde{\theta}_{ij}(t) = (-\alpha)^{-j+i-1} \left[e^{-\alpha t} \left(e^{\alpha t} - \sum_{l=j-i+1}^{\infty} \frac{(\alpha t)^l}{l!} \right) - 1 \right],$$

autrement dit

$$\tilde{\theta}_{ij}(t) = \frac{(-1)^{j-i}}{\alpha^{j-i+1}} e^{-\alpha t} \left(\sum_{l=j-i+1}^{\infty} \frac{(\alpha t)^l}{l!} \right). \quad (4.11)$$

ii) Le produit matriciel $e^{Jt} \int_0^t e^{-Js} ds$, noté

$$[W_{ij}(t)] = [\theta_{ij}(t)] [\tilde{\theta}_{ij}(t)],$$

vérifie

$$\begin{aligned}W_{ij}(t) &= \sum_{n=1}^m \theta_{in}(t) \tilde{\theta}_{nj}(t) \\ &= \sum_{n=i}^m \frac{t^{n-i}}{(n-i)!} e^{\alpha t} \frac{(-1)^{j-n}}{\alpha^{j-n+1}} e^{-\alpha t} \left(\sum_{l=j-n+1}^{\infty} \frac{(\alpha t)^l}{l!} \right) \\ &= \sum_{n=i}^{m_L} \sum_{l=j-n+1}^{\infty} (-1)^{j-n} \frac{\alpha^l t^{n-i+l}}{l!(n-i)! \alpha^{j-n+1}} \\ &= \sum_{n=i}^m \sum_{l=j-n+1}^{\infty} \frac{(-1)^{j-n} \alpha^{l+n-1-j} t^{l+n-i}}{l!(n-i)!}.\end{aligned}$$

Si on pose $n' = n - i$ et $l' = l - j + n - 1$, l'équation devient :

$$W_{ij}(t) = (-t)^{j-i} \sum_{l'=0}^{\infty} \sum_{n'=0}^{m-i} \frac{(-1)^{n'} \alpha^{l'} t^{l'+1}}{n'!(l' + j - i + 1 - n')!},$$

d'où

$$W_{ij}(t) = (-1)^{j-i} t^{j-i+1} \sum_{l'=0}^{\infty} (\alpha t)^{l'} \sum_{n'=0}^{m-i} \frac{(-1)^{n'}}{n'!(l' + j - i + 1 - n')!}. \quad (4.12)$$

Autre écriture :

L'expression du produit matriciel $e^{Jt} \int_0^t e^{-Js} ds$ peut s'exprimer de façon plus compacte de la façon suivante :

Puisque $e^{Jt} = \sum_{k=0}^{\infty} \frac{(tJ)^k}{k!}$, on obtient que

$$\begin{aligned} \int_0^t e^{-Js} ds &= \int_0^t \left(\mathbb{I} + \frac{(-s)J}{1!} + \frac{(-s)^2 J^2}{2!} + \dots + \frac{(-s)^{m-1} J^{m-1}}{(m-1)!} + \dots \right) ds \\ &= t \left(\sum_{k=0}^{\infty} (-1)^{k'} \frac{(tJ)^{k'}}{(k'+1)!} \right) \end{aligned}$$

donc

$$\begin{aligned} e^{Jt} \int_0^t e^{-Js} ds &= t \left(\sum_{k=0}^{\infty} \frac{(tJ)^k}{k!} \right) \left(\sum_{k=0}^{\infty} (-1)^{k'} \frac{(tJ)^{k'}}{(k'+1)!} \right) \\ &= t \sum_{k,k'=0}^{\infty} (-1)^{k'} \frac{(tJ)^{k+k'}}{k!(k'+1)!}. \end{aligned}$$

Par ailleurs, $J = \alpha \mathbb{I} + \mathbf{E}$ vérifie $J^k = \sum_{l=0}^k C_k^l \alpha^l \mathbf{E}^{k-l}$, ce qui donne

$$e^{Jt} \int_0^t e^{-Js} ds = t \sum_{k,k'=0}^{\infty} \frac{(-1)^{k'} t^{k+k'}}{k!(k'+1)!} \left(\sum_{l'=0}^{k+k'} C_{k+k'}^{l'} \alpha^{l'} \mathbf{E}^{k+k'-l'} \right).$$

D'un autre côté, on sait que :

$$0 = (1-1)^{l+1} = \sum_{k=0}^{l+1} C_{l+1}^k 1^k (-1)^{l-k+1} = (l+1)! \sum_{k=0}^l \frac{(-1)^{l-k+1}}{k!(l-k+1)!} + 1,$$

D'où $\sum_{k=0}^l \frac{(-1)^{k'+1}}{k!(k'+1)!} = -\frac{1}{(l+1)!}$, en posant $l = k + k'$.

Au final, en intervertissant les sommes, on a :

$$e^{Jt} \int_0^t e^{-Js} ds = t \sum_{l=0}^{\infty} t^l \sum_{l'=0}^l C_l^{l'} \alpha^{l'} \mathbf{E}^{l-l'} \sum_{k,k'=0}^l \frac{(-1)^{k'}}{k!(k'+1)!} = t \sum_{l=0}^{\infty} t^l \sum_{l'=0}^l \frac{C_l^{l'} \alpha^{l'} \mathbf{E}^{l-l'}}{(l+1)!}.$$

autrement dit,

$$e^{Jt} \int_0^t e^{-Js} ds = t \sum_{l=0}^{\infty} t^l \frac{J^l}{(l+1)!}. \quad (4.13)$$

Ainsi, pour une valeur propre α ,

— si $\alpha \neq 0$, la matrice J est inversible, et $Je^{Jt} \int_0^t e^{-Js} ds = e^{Jt} - \mathbb{I}$, c.à.d.

$$e^{Jt} \int_0^t e^{-Js} ds = J^{-1}(e^{Jt} - \mathbb{I}). \quad (4.14)$$

— si $\alpha = 0$, on aura $J^l = 0$ si $l > 1$, d'où

$$e^{Jt} \int_0^t e^{-Js} ds = \sum_{l=0}^{m-1} \frac{t^{l+1}}{(l+1)!} \mathbf{E}^l. \quad (4.15)$$

Remarque :

Pour une valeur propre complexe, $\alpha = \alpha_1 + i\alpha_2$, on aura :

$$e^{-J s} = e^{-\alpha_1 s} (\cos(\alpha_2 s) + i \sin(\alpha_2 s)) \sum_{l=0}^{m-1} \frac{(-s)^l \mathbf{E}^l}{l!}.$$

On voit que matrice exponentielle, en fonction de α_1 , peut amener à des comportements périodiques explosifs ou périodiques amorties vers l'extinction, pour telle ou telle composante du processus.

Recherche de la solution générale

I) On a ainsi trouvé, pour tous $\mathbf{y} = (y_1, \dots, y_K)$ et $\nu = -(\mu_1, \dots, \mu_K)$, que

$$\mathcal{I}(t, \mathbf{y}) = e^{At} \left[\mathbf{y} + \left(\int_0^t e^{-As} ds \right) \nu \right] \quad \text{avec} \quad \mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1, \dots, \mathcal{I}_K)(t, \mathbf{y}).$$

Par conséquent, on a :

$$\mathcal{I}(t, \mathbf{y}) = \mathbf{z} \quad \Rightarrow \quad \mathbf{y} + \int_0^t e^{-As} ds \nu = e^{-At} \mathcal{I}(t, \mathbf{y}) = e^{-At} \mathbf{z},$$

d'où $\mathbf{y} = e^{-At} \mathbf{z} - \int_0^t e^{-As} ds \nu$. Autrement dit, si on note $e^{-As} = (\theta_{ij}(t))$,

$$\mathbf{y} = \mathbf{G}(t, \mathbf{z}) = (\theta_{ij}(t))^{-1} \mathbf{z} + \tilde{\theta}_{ij}(t) \nu.$$

Donc

$$G_i(t, \mathbf{z}) = \sum_{j=1}^K \beta_{ij}(t) z_j + \beta_{i0}(t), \quad \beta_{i0} = \tilde{\theta}_{ij} \nu, \quad i = 1, \dots, K. \quad (4.16)$$

Ainsi chaque G_i est une multinomiale de dimension inférieure ou égale à K , alourdie en 0 et de taille totale (somme des coordonnées) ≤ 1 .

- $\beta_{i0}(t) = \mathbb{P}(\mathbf{X}(t) = 0 \mid \mathbf{X}(0) = e_i)$,
- $\beta_{ij}(t) = \mathbb{P}(\mathbf{X}(t) = e_j \mid \mathbf{X}(0) = e_i)$.

II) Ainsi l'équation finale de $\mathcal{I}_0(t, \tilde{\mathbf{y}})$ est

$$\mathcal{I}'_0(t, \tilde{\mathbf{y}}) = \Lambda_0 \mathcal{I}_0(t, \tilde{\mathbf{y}}) - \lambda_{00} \mathcal{I}_0^2(t, \tilde{\mathbf{y}}) - \lambda_0^T \mathcal{I}(t, \mathbf{y}) - \mu_0 \quad (4.17)$$

$$\text{avec } \begin{cases} \lambda_0^T = (\lambda_{01}, \dots, \lambda_{0K}) \\ \tilde{\mathbf{y}} = (y_0, \mathbf{y}) \end{cases} \quad \text{et la condition initiale } \mathcal{I}_0(0, \tilde{\mathbf{y}}) = y_0.$$

Résolution de (4.17) :

Supposons que l'on sache calculer une solution particulière $J_0(t)$ satisfaisant (4.17)

a) Soit $U(t, \tilde{\mathbf{y}}) = \mathcal{I}_0(t, \tilde{\mathbf{y}}) - J_0(t)$ et soit $\Theta(t, \mathbf{y}) = \lambda_0^T \mathcal{I}(t, \mathbf{y}) + \mu_0$, alors

$$\begin{aligned} U'(t, \tilde{\mathbf{y}}) &= \mathcal{I}'_0(t, \tilde{\mathbf{y}}) - J'_0(t) \\ &= \Lambda_0 \mathcal{I}_0(t, \tilde{\mathbf{y}}) - \lambda_{00} \mathcal{I}_0^2(t, \tilde{\mathbf{y}}) - \Theta(t, \mathbf{y}) - [\Lambda_0 J_0(t) - \lambda_{00} J_0^2(t) - \Theta(t, \mathbf{y})] \\ &= \Lambda_0 (\mathcal{I}_0(t, \tilde{\mathbf{y}}) - J_0(t)) - \lambda_{00} (\mathcal{I}_0^2(t, \tilde{\mathbf{y}}) - J_0^2(t)) \\ &= (\Lambda_0 - 2\lambda_{00} J_0(t)) (\mathcal{I}_0(t, \tilde{\mathbf{y}}) - J_0(t)) - \lambda_{00} (\mathcal{I}_0(t, \tilde{\mathbf{y}}) - J_0(t))^2, \end{aligned}$$

autrement dit,

$$U'(t, \tilde{\mathbf{y}}) = (\Lambda_0 - 2\lambda_{00} J_0(t)) U(t, \tilde{\mathbf{y}}) - \lambda_{00} U^2(t, \tilde{\mathbf{y}}). \quad (4.18)$$

b) Par ailleurs, $D(t, \tilde{\mathbf{y}}) = \frac{1}{U(t, \tilde{\mathbf{y}})}$ vérifie $D'(t, \tilde{\mathbf{y}}) = -\frac{U'}{U^2}(t, \tilde{\mathbf{y}})$, c.à.d.

$$\frac{U'}{U^2}(t, \tilde{\mathbf{y}}) = (\Lambda_0 - 2\lambda_{00} J_0(t)) \frac{1}{U(t, \tilde{\mathbf{y}})} - \lambda_{00}.$$

Ainsi,

$$D'(t, \tilde{\mathbf{y}}) = -(\Lambda_0 - 2\lambda_{00} J_0(t)) D(t, \tilde{\mathbf{y}}) + \lambda_{00}. \quad (4.19)$$

Sa solution vaut

$$D(t, \tilde{\mathbf{y}}) = e^{-\int_0^t (\Lambda_0 - 2\lambda_{00} J_0(s)) ds} \left[D(0, \tilde{\mathbf{y}}) + \lambda_{00} \int_0^t e^{\int_0^s (\Lambda_0 - 2\lambda_{00} J_0(u)) du} ds \right]. \quad (4.20)$$

Soient

$$\begin{cases} n(t) = \Lambda_0 - 2\lambda_{00} J_0(t), \\ N(t) = \int_0^t n(s) ds = \Lambda_0 t - 2\lambda_{00} \int_0^t J_0(s) ds, \\ \Gamma(t) = e^{N(t)}. \end{cases}$$

On a alors :

$$\begin{aligned} D(t, \tilde{\mathbf{y}}) &= \frac{1}{\mathcal{I}_0(t, \tilde{\mathbf{y}}) - J_0(t)} \Rightarrow \mathcal{I}_0(t, \tilde{\mathbf{y}}) = J_0(t) + \frac{1}{D(t, \tilde{\mathbf{y}})} \\ \mathcal{I}_0(t, \tilde{\mathbf{y}}) &= J_0(t) + \frac{\Gamma(t)}{D(0, \tilde{\mathbf{y}}) + \lambda_{00} \int_0^t \Gamma(s) ds}. \end{aligned} \quad (4.21)$$

c) Au final, si $\mathcal{I}_0(t, \tilde{\mathbf{y}}) = z_0$, on aura $z_0 = \mathcal{I}_0(t, y_0, G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z}))$ car

$$\begin{cases} \mathcal{I}_1(t, \mathbf{y}) = z_1 \\ \vdots \\ \mathcal{I}_K(t, \mathbf{y}) = z_K \end{cases} \Rightarrow \begin{cases} y_1 = G_1(t, \mathbf{z}) \\ \vdots \\ y_K = G_K(t, \mathbf{z}). \end{cases}$$

$G_0(t, \mathbf{z}) = y_0$ sera solution de l'équation implicite :

$$\text{c.à.d.} \quad \begin{cases} \mathcal{I}_0(t, y_0, G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})) = z_0 \\ G_0(t, \tilde{\mathbf{z}}) = y_0(t, \tilde{\mathbf{z}}). \end{cases} \quad (4.22)$$

Méthode générale de la recherche des solutions :

On sait que la solution de l'équation matricielle

$$X'(t) = B(t) + A(t)X(t) - X(t)D(t) - X(t)C(t)X(t)$$

est de la forme $X(t) = U(t)F^{-1}(t)$, où U et F sont solutions du système homogène

$$\begin{pmatrix} U'(t) \\ F'(t) \end{pmatrix} = \begin{pmatrix} A(t) & B(t) \\ C(t) & D(t) \end{pmatrix} \begin{pmatrix} U(t) \\ F(t) \end{pmatrix},$$

si F est inversible.

On réécrit notre équation de branchement (4.17) sous la forme suivante :

$$\mathcal{I}'_0(t, \tilde{\mathbf{y}}) = \gamma(t) + \alpha \mathcal{I}_0(t, \tilde{\mathbf{y}}) + \beta \mathcal{I}_0^2(t, \tilde{\mathbf{y}}). \quad (4.23)$$

Elle est scalaire et correspond aux paramètres :

$$\begin{cases} B(t) = \gamma(t) = -H(t, \mathbf{y}), \\ A(t) = \alpha = \Lambda_0, \\ C(t) = -\beta = \lambda_{00}, \\ D(t) = 0. \end{cases}$$

On l'associe alors à l'équation linéaire (4.23)

$$\begin{pmatrix} U'(t) \\ F'(t) \end{pmatrix} = \begin{pmatrix} \alpha & \gamma(t) \\ -\beta & 0 \end{pmatrix} \begin{pmatrix} U(t) \\ F(t) \end{pmatrix}. \quad (4.24)$$

Cela correspond aux équations :
$$\begin{cases} F'(t) = -\beta U(t), \\ U'(t) = \alpha U(t) + \gamma(t)F(t). \end{cases}$$

Ainsi la relation $F(t) = \gamma_0 - \beta \int_0^t U(s)ds$, permet d'écrire la solution générale de (4.23) sous la forme :

$$\mathcal{I}_0(t, \tilde{\mathbf{y}}) = \frac{U(t)}{\gamma_0 - \beta \int_0^t U(s)ds} \quad \text{ou encore mieux} \quad \mathcal{I}_0(t, \tilde{\mathbf{y}}) = \frac{F'}{-\beta F} \quad (4.25)$$

avec $F(t)$ vérifiant l'EDO homogène du second d'ordre :

$$F''(t) - \alpha F'(t) + \beta\gamma(t)F(t) = 0. \quad (4.26)$$

Ainsi on a :

$$-\beta \int_0^t \mathcal{I}_0(s, \tilde{\mathbf{y}}) ds = \log \left(\frac{F(t)}{F(0)} \right) \Rightarrow F(t) = \exp \left(-\beta \int_0^t \mathcal{I}_0(s, \tilde{\mathbf{y}}) ds + cste \right).$$

On peut réécrire l'équation (4.24), en prenant $\begin{cases} \tilde{u} = F \\ \tilde{v} = F' \end{cases}$

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}' = \begin{pmatrix} F' \\ F'' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \gamma(t) & \alpha \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \quad (4.27)$$

Pour trouver la solution générale $\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}$ de l'EDO homogène (4.27), il suffit de connaître deux solutions fondamentales $u_1(t)$ et $u_2(t)$ puis de déterminer la forme de la résolvante :

$$R(t, 0) = \begin{pmatrix} u_1 & u_2 \\ u_1' & u_2' \end{pmatrix} (t) \quad \text{avec} \quad R(0, 0) = \mathbb{I}_2$$

avec les conditions initiales $\begin{cases} u_1(0) = 1, & u_1'(0) = 0, \\ u_2(0) = 0, & u_2'(0) = 1. \end{cases}$

La solution générale de l'EDO (4.26) peut s'écrire alors sous la forme suivante :

$$F(t) = \rho_1 u_1(t) + \rho_2 u_2(t),$$

avec

$$\begin{cases} F(0) = \rho_1 = \gamma_0, \\ F'(0) = \rho_2 = -\beta\gamma_0 y_0. \end{cases}$$

En réalité, on a besoin de connaître une solution particulière en utilisant le Wronskien. En effet, le Wronskien $W(t) = u_1(t)u_2'(t) - u_2(t)u_1'(t)$ a pour dérivée

$$W'(t) = u_1(t)u_2''(t) - u_2(t)u_1''(t),$$

puisque $u_1(t)$ et $u_2(t)$ sont des solutions fondamentales de l'équation homogène (4.27). On s'aperçoit que

$$W'(t) = \alpha(u_1(t)u_2'(t) - u_2(t)u_1'(t)) = \alpha W(t)$$

c.à.d. $W(t) = e^{\alpha t}$ car $W(0) = 1$.

Ainsi, si u_1 est connue, on aura $W_0 e^{\alpha t} = u_1(t)u_2'(t) - u_2(t)u_1'(t)$ i.e.

$$u_2'(t) = u_2(t) \frac{u_1'(t)}{u_1(t)} + \frac{e^{\alpha t}}{u_1(t)}$$

de solution

$$u_2(t) = u_1(t) \int_0^t \frac{e^{\alpha s}}{u_1^2(s)} ds. \quad (4.28)$$

Il suffit donc de trouver une solution fondamentale $u_1(t)$ de l'EDO homogène suivante :

$$u_1''(t) - \alpha u_1'(t) + \beta\gamma(t)u_1(t) = 0 \quad (4.29)$$

On peut normaliser cette équation en une forme standard :

$$\begin{cases} u_1(t) = x(t)e^{\theta(t)} \\ u_1'(t) = x'(t)e^{\theta(t)} + x(t)\theta'(t)e^{\theta(t)} \\ u_1''(t) = e^{\theta(t)} [x''(t) + 2x'(t)\theta'(t) + x(t)((\theta'(t))^2 + \theta''(t))] \end{cases}$$

en remplaçant $u_1(t), u_1'(t), u_1''(t)$ dans (4.29), on aura l'EDO homogène suivante :

$$x''(t) + x'(t)(2\theta'(t) - \alpha) + (\theta''(t) + (\theta'(t))^2 - \alpha\theta'(t) + \beta\gamma(t))x(t) = 0$$

on impose $2\theta'(t) - \alpha = 0$, donc $\theta(t) = \frac{\alpha}{2}t$, ce qui donne l'EDO homogène standardisée suivante :

$$x''(t) + \left(\beta\gamma(t) - \frac{\alpha^2}{4} \right) x(t) = 0. \quad (4.30)$$

Dans notre cas, $\gamma(t) = \sum_{j=0}^{K-1} A_j(t)e^{\Lambda_j t}$, avec $A_j(t)$ sont des fonctions polynomiales de t de degré $\leq K - 1$.

Des exemples particuliers avec $m = 3$ et 4 illustreront cette approche.

Par ailleurs il existe des méthodes numériques pour la recherche de telles solutions particulières (voir les figures 4.5,4.6,4.7)

4.4.4 Résolution numérique des EDO

1. On donne l'exemple de résolution numérique de l'équation (4.30) de $x(t)$, avec $K = 4$ populations et la condition initiale $x(0) = 1$, $x'(0) = -\frac{1}{2}$. Ici, on a choisi les paramètres suivants : $\beta = 0.1$, $\alpha = 2$ et $\gamma(t) = \sum_{j=0}^3 A_j(t)e^{\Lambda_j t}$ avec $A_1 = 6$, $A_2 = 10$, $A_3 = 4$, $\Lambda_1 = 0.5$, $\Lambda_2 = 0.2$ et $\Lambda_3 = 0.9$.

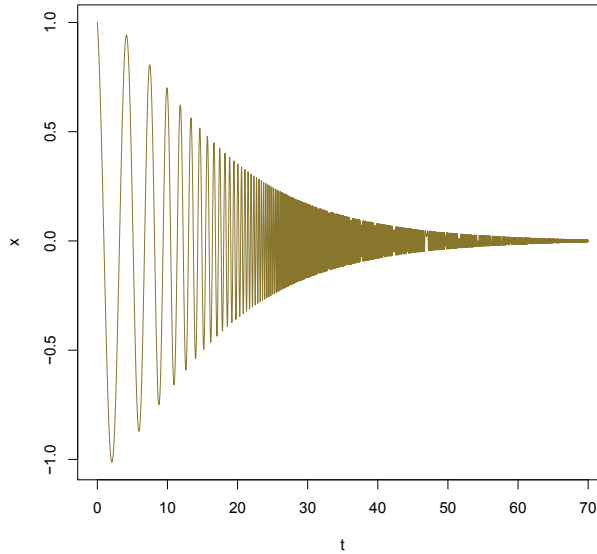


FIGURE 4.5 – Graphe de la solution $x(t)$ de l'EDO homogène du second ordre (4.30) s'amortissant dans le temps vers 0.

2. Le graphe 4.6 représente les deux solutions fondamentales de l'équation homogène (4.26) par rapport au temps. Elles sont déduites de $x(t)$: $u_1(t)$ (à gauche) de l'EDO (4.29) et ensuite $u_2(t)$ (à droite) de l'équation (4.28), vérifiant les conditions initiales

$$\begin{cases} u_1(0) = 1, & u_1'(0) = 0 \\ u_2(0) = 0, & u_2'(0) = 1. \end{cases}$$

Pour la première solution fondamentale $u_1(t)$, le paramètre $\alpha = 2$ donne $\theta(t) = \frac{\alpha}{2}t$. Ainsi on a $u_1(t) = x(t)e^{\theta(t)}$. Ensuite, de l'équation (4.28), on déduit la solution fondamentale $u_2(t)$.

Remarque :

Les calculs ont été réalisés sous logiciel R avec le package “desolve”. Dans l'algorithme, on a utilisé la méthode des rectangles à droite pour l'approximation numérique de l'intégrale $\int_0^t \frac{e^{\alpha s}}{u_1^2(s)} ds$.

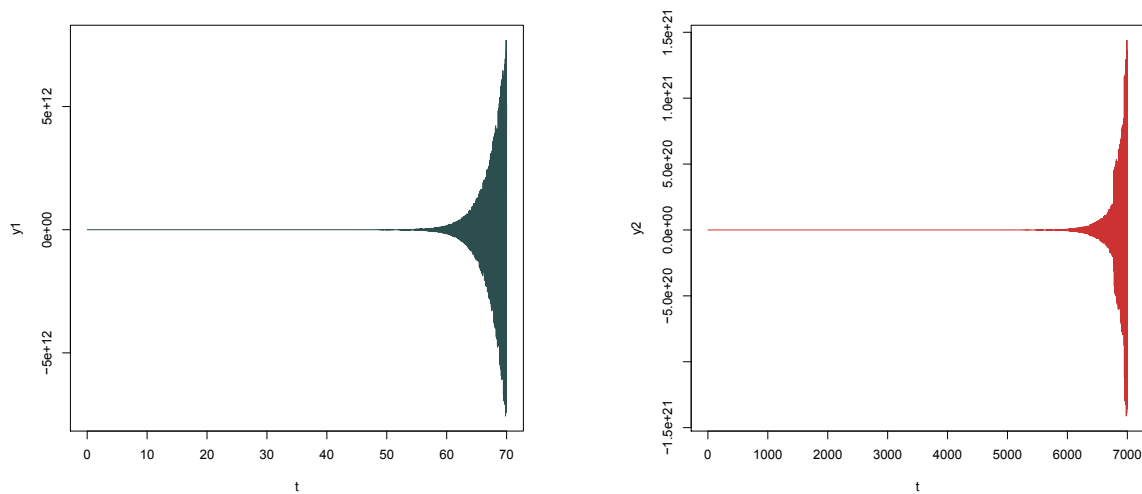


FIGURE 4.6 – Représentation numérique des solutions fondamentales $u_1(t)$ et $u_2(t)$.

3. Le graphe 4.7 représente la résolution numérique de $\mathcal{I}_0(t, \tilde{\mathbf{y}}) = -\frac{F'(t)}{\beta F(t)}$. La condition initiale $\mathcal{I}_0(0, \tilde{\mathbf{y}}) = y_0 = 10$ implique le choix des conditions initiales pour $F(t)$ avec $F(0) = \rho_1 = \gamma_0 = 0.3$ et $F'(0) = \rho_2 = -\beta\gamma_0 y_0 = -0.3$. On a utilisé les solutions fondamentales numériques précédentes, à savoir $u_1(t)$ et $u_2(t)$, pour définir $F(t) = \rho_1 u_1(t) + \rho_2 u_2(t)$.

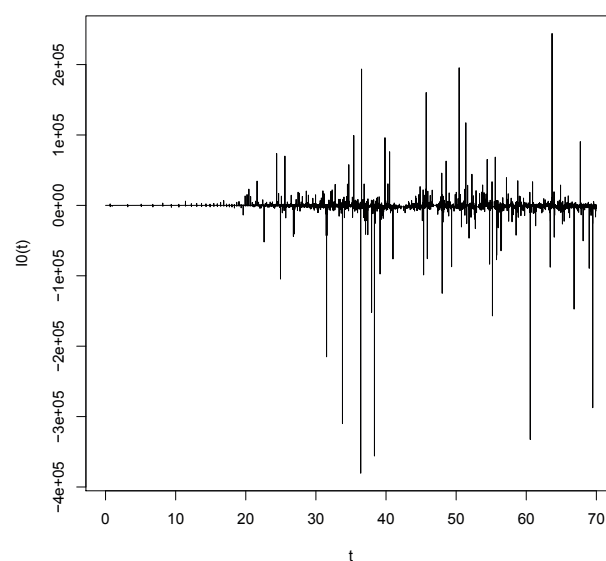


FIGURE 4.7 – Représentation numérique de la solution $\mathcal{I}_0(t, \tilde{\mathbf{y}})$ (4.25) de la classe “0”.

4.5 Un exemple de branchement 3-types

Nous considérons un modèle à $K = 3$ populations du type quadratique précédent, Ainsi pour la classe "0" on a :

$$V_0(z_0, \mathbf{z}) = \mu_0 - \Lambda_0 z_0 + \lambda_{00} z_0^2 + \lambda_{01} z_1 + \lambda_{02} z_2, \quad \Lambda_0 = \mu_0 + \lambda_{00} + \lambda_{01} + \lambda_{02},$$

et pour les deux autres classes :

$$\begin{cases} V_1(\mathbf{z}) = \mu_1 - \Lambda_1 z_1 + \lambda_{12} z_2 \\ V_2(\mathbf{z}) = \mu_2 - \Lambda_2 z_2 + \lambda_{21} z_1 \end{cases}, \quad \mathbf{z} = (z_1, z_2),$$

avec

$$\begin{cases} \Lambda_1 = \mu_1 + \lambda_{12} \\ \Lambda_2 = \mu_2 + \lambda_{21} \end{cases}.$$

i) Soit donc $\mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1(t, \mathbf{y}), \mathcal{I}_2(t, \mathbf{y}))$ satisfait l'EDO :

$$\frac{d\mathcal{I}}{dt}(t, \mathbf{y}) = -\mathbf{V}(\mathcal{I}(t, \mathbf{y}))$$

c.à.d.

$$\frac{d\mathcal{I}}{dt} = A\mathcal{I} + \nu$$

avec

$$A = \begin{pmatrix} \Lambda_1 & -\lambda_{12} \\ -\lambda_{21} & \Lambda_2 \end{pmatrix},$$

et

$$-\nu^T = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 & -\lambda_{12} \\ -\lambda_{21} & \Lambda_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = A\mathbf{1}.$$

Si $\mathcal{I}^T(0) = \mathbf{y}$, sa solution est donnée par ce qui suit :

$$\begin{aligned} \mathcal{I}(t, \mathbf{y}) &= e^{At} \left[\mathbf{y} + \left(\int_0^t e^{-As} ds \right) \nu \right] \\ &= e^{At} \left[\mathbf{y} - \left(\int_0^t e^{-As} ds \right) (A\mathbf{1}) \right] \\ &= e^{At} \left[\mathbf{y} + (e^{-At} - \mathbf{1}) \mathbf{1} \right] \\ &= e^{At} (\mathbf{y} - \mathbf{1}) + \mathbf{1}. \end{aligned}$$

ii) Une matrice de passage P (ou P^{-1}) de Jordan pour A est facile à établir. En effet, commençons par la recherche des valeurs propres :

$$|A - \alpha \mathcal{I}| = \alpha^2 - (\Lambda_1 + \Lambda_2)\alpha + \Lambda_1 \Lambda_2 - \lambda_{12} \lambda_{21}.$$

On voit que son discriminant $\Delta = (\Lambda_1 - \Lambda_2)^2 + \lambda_{12} \lambda_{21} \geq 0$. On aura ainsi deux cas à traiter :

Cas 1 : $\Delta > 0$: 2 valeurs propres distinctes.

Cas 2 : $\Delta = 0$: 1 valeur propre double qui ne survient que dans les cas suivants :

$$\begin{cases} \text{Cas 2.1 : } \lambda_{12} = 0 \text{ avec } \lambda_{21} \neq 0 \text{ et } \mu_1 = \mu_2 + \lambda_{21}, \\ \text{Cas 2.2 : } \lambda_{21} = 0 \text{ avec } \lambda_{12} \neq 0 \text{ et } \mu_2 = \mu_1 + \lambda_{12}, \\ \text{Cas 2.3 : } \lambda_{12} = \lambda_{21} = 0 \text{ et } \mu_1 = \mu_2. \end{cases} \quad (4.31)$$

Cas 1 : Valeurs propres distinctes

Dans ce cas, $\alpha_{1,2} = \frac{1}{2}(\Lambda_1 + \Lambda_2 \pm \sqrt{\Delta})$. On peut prendre comme matrice de passage

$$P = \begin{pmatrix} \frac{\Lambda_2 - \alpha_1}{\lambda_{21}} & \frac{\Lambda_2 - \alpha_2}{\lambda_{21}} \\ 1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} \frac{\lambda_{21}}{\sqrt{\Delta}} & \frac{\alpha_2 - \Lambda_2}{\lambda_{21}} \\ \frac{-\lambda_{21}}{\sqrt{\Delta}} & \frac{\Lambda_2 - \alpha_1}{\lambda_{21}} \end{pmatrix},$$

de cette façon $J = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$, $\tilde{\nu} = P\nu = - \begin{pmatrix} \left(\frac{\Lambda_2 - \alpha_1}{\lambda_{21}}\right) \mu_1 + \left(\frac{\Lambda_2 - \alpha_2}{\lambda_{21}}\right) \mu_2 \\ \mu_1 + \mu_2 \end{pmatrix}$,

et $\tilde{\mathbf{y}} = P\mathbf{y} = - \begin{pmatrix} \left(\frac{\Lambda_2 - \alpha_1}{\lambda_{21}}\right) y_1 + \left(\frac{\Lambda_2 - \alpha_2}{\lambda_{21}}\right) y_2 \\ y_1 + y_2 \end{pmatrix}$.

La décomposition de A en blocs de Jordan donne $e^{-As} = P^{-1}e^{-Js}P$.

En prenant $\tilde{\mathcal{I}}(t, \mathbf{y}) = P\mathcal{I}(t, \mathbf{y})$ et $\tilde{\nu} = P\nu$, on obtient la solution :

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = e^{Jt} \left(\tilde{\mathcal{I}}(0, \mathbf{y}) + \int_0^t e^{-Js} \tilde{\nu} ds \right),$$

où $\int_0^t e^{-Js} ds = \begin{pmatrix} \frac{(1-e^{-\alpha_1 t})}{\alpha_1} & 0 \\ 0 & \frac{(1-e^{-\alpha_2 t})}{\alpha_2} \end{pmatrix}$,

c'est à dire :

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} \frac{e^{\alpha_1 t}}{\lambda_{21}} ((\Lambda_2 - \alpha_1)y_1 + (\Lambda_2 - \alpha_2)y_2) - \frac{1-e^{\alpha_1 t}}{\alpha_1 \lambda_{21}} ((\Lambda_2 - \alpha_1)\mu_1 + (\Lambda_2 - \alpha_2)\mu_2) \\ e^{\alpha_2 t} (y_1 + y_2) + \frac{1-e^{\alpha_2 t}}{\alpha_2} (\mu_1 + \mu_2) \end{pmatrix}.$$

La solution $\mathcal{I}(t, \mathbf{y})$, partant de \mathbf{y} , est donnée par $\mathcal{I}(t, \mathbf{y}) = P^{-1}\tilde{\mathcal{I}}(t, \mathbf{y})$

$$\mathcal{I}(t, \mathbf{y}) = \begin{cases} \mathcal{I}_1(t, \mathbf{y}) = \theta_{10}(t) + \theta_{11}(t)y_1 + \theta_{12}(t)y_2 \\ \mathcal{I}_2(t, \mathbf{y}) = \theta_{20}(t) + \theta_{21}(t)y_1 + \theta_{22}(t)y_2 \end{cases},$$

autrement dit :

$$\mathcal{I}(t, \mathbf{y}) = \begin{pmatrix} \theta_{11}(t) & \theta_{12}(t) \\ \theta_{21}(t) & \theta_{22}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \theta_{10}(t) \\ \theta_{20}(t) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

D'où

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_{11}(t) & \theta_{12}(t) \\ \theta_{21}(t) & \theta_{22}(t) \end{pmatrix}^{-1} \begin{pmatrix} z_1 - \theta_{10}(t) \\ z_2 - \theta_{20}(t) \end{pmatrix}.$$

Donc, pour $\tilde{\mathbf{z}} = (z_0, z_1, z_2)$, on a

$$\begin{aligned} G_1(t, \tilde{\mathbf{z}}) &= \beta_{10}(t) + 0z_0 + \beta_{11}(t)z_1 + \beta_{12}(t)z_2, \\ G_2(t, \tilde{\mathbf{z}}) &= \beta_{20}(t) + 0z_0 + \beta_{21}(t)z_1 + \beta_{22}(t)z_2. \end{aligned}$$

Conclusion :

Pour tous les $i \geq 1$, $G_i(t, \tilde{\mathbf{z}})$ est une loi multinomiale de

$\mathbf{X}(t) = (X_0(t), X_1(t), X_2(t))$ de masse totale 1, alourdie en "0" = (0, 0, 0),

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0)) = \beta_{i0}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 1, 0)) = \beta_{i1}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 1)) = \beta_{i2}(t). \end{cases}$$

Cas 2 : Une valeur propre double

1. Prenant le cas 2.1 de (4.31) $\lambda_{12} = 0$ avec $\lambda_{21} \neq 0$ et $\mu_1 = \mu_2 + \lambda_{21}$, *i.e.*

$$A = \begin{pmatrix} \Lambda_1 & 0 \\ -\lambda_{21} & \Lambda_1 \end{pmatrix}.$$

On a comme matrice de passage

$$P = \begin{pmatrix} 0 & -\frac{1}{\lambda_{21}} \\ 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 1 \\ -\lambda_{21} & 0 \end{pmatrix}.$$

De cette façon, on obtient

$$J = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}, \quad \tilde{\nu} = P\nu = \begin{pmatrix} \frac{\mu_2}{\lambda_{21}} \\ -\mu_1 \end{pmatrix} \text{ et } \tilde{\mathbf{y}} = P\mathbf{y} = - \begin{pmatrix} \frac{-y_2}{\lambda_{21}} \\ y_1 \end{pmatrix},$$

puis :

$$\int_0^t e^{-Js} ds = \begin{pmatrix} \frac{(1-e^{-\alpha t})}{\Lambda_1} & \frac{-t^2}{2} \\ 0 & \frac{(1-e^{-\alpha t})}{\Lambda_1} \end{pmatrix},$$

et la solution est

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} (e^{-\Lambda_1 t} - 1) \left(\mu_1 t - \frac{\mu_2 e^{\Lambda_1 t}}{\lambda_{21}} \right) + \frac{\mu_1 t^2 e^{\Lambda_1 t}}{2} + y_1 t - \frac{y_2 e^{\Lambda_1 t}}{\lambda_{21}} \\ y_1 e^{\Lambda_1 t} + \frac{\mu_1}{\Lambda_1} (1 - e^{\Lambda_1 t}) \end{pmatrix}.$$

Au final, $\mathcal{I}(t)$ en fonction de \mathbf{y} est donnée par :

$$\mathcal{I}(t, \mathbf{y}) = P^{-1}\tilde{\mathcal{I}}(t) = \begin{pmatrix} y_1 e^{\Lambda_1 t} + \frac{\mu_1}{\Lambda_1} (1 - e^{\Lambda_1 t}) \\ e^{\Lambda_1 t} y_2 - \lambda_{21} y_1 t + \frac{\mu_2}{\Lambda_1} (1 - e^{\Lambda_1 t}) - \frac{\lambda_{21} \mu_1 t}{\Lambda_1} (e^{-\Lambda_1 t} - 1) - \frac{\lambda_{21} \mu_1 t^2}{2} e^{\Lambda_1 t} \end{pmatrix}.$$

Autrement dit :

$$\begin{aligned} \mathcal{I}(t, \mathbf{y}) &= \begin{pmatrix} e^{\Lambda_1 t} & 0 \\ -\lambda_{21} t & e^{\Lambda_1 t} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \frac{\mu_1}{\Lambda_1} (1 - e^{\Lambda_1 t}) \\ \frac{1 - e^{\Lambda_1 t}}{\Lambda_1} (\mu_2 - \lambda_{21} \mu_1 t e^{-\Lambda_1 t}) - \frac{\lambda_{21} \mu_1 t^2}{2} e^{\Lambda_1 t} \end{pmatrix} \\ &= \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \end{aligned}$$

d'où

$$\begin{aligned} \begin{pmatrix} G_1(t, \tilde{\mathbf{z}}) \\ G_2(t, \tilde{\mathbf{z}}) \end{pmatrix} &= \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \begin{pmatrix} e^{\Lambda_1 t} & 0 \\ -\lambda_{21} t & e^{\Lambda_1 t} \end{pmatrix}^{-1} \begin{pmatrix} z_1 - \frac{\mu_1}{\Lambda_1} (1 - e^{\Lambda_1 t}) \\ z_2 - \frac{1 - e^{\Lambda_1 t}}{\Lambda_1} (\mu_2 - \lambda_{21} \mu_1 t e^{-\Lambda_1 t}) - \frac{\lambda_{21} \mu_1 t^2}{2} e^{\Lambda_1 t} \end{pmatrix} \\ &= \begin{pmatrix} e^{-\Lambda_1 t} & 0 \\ \lambda_{21} t e^{-2\Lambda_1 t} & e^{-\Lambda_1 t} \end{pmatrix} \begin{pmatrix} z_1 - \frac{\mu_1}{\Lambda_1} (1 - e^{\Lambda_1 t}) \\ z_2 - \frac{1 - e^{\Lambda_1 t}}{\Lambda_1} (\mu_2 - \lambda_{21} \mu_1 t e^{-\Lambda_1 t}) - \frac{\lambda_{21} \mu_1 t^2}{2} e^{\Lambda_1 t} \end{pmatrix}. \end{aligned}$$

Donc

$$\begin{pmatrix} G_1(t, \tilde{\mathbf{z}}) \\ G_2(t, \tilde{\mathbf{z}}) \end{pmatrix} = \begin{pmatrix} e^{-\Lambda_1 t} z_1 - \frac{\mu_1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) \\ \lambda_{21} t e^{-2\Lambda_1 t} z_1 + e^{-\Lambda_1 t} z_2 - \frac{e^{-\Lambda_1 t} - 1}{\Lambda_1} (\lambda_{21} e^{-\Lambda_1 t} (\Lambda_1 t - \mu_1) + \mu_2) - \frac{\lambda_{21} \mu_1 t^2}{2} \end{pmatrix}.$$

Conclusion :

Pour tout $i \geq 1$, $G_i(t, \tilde{\mathbf{z}})$ est une loi multinomiale de $\mathbf{X}(t) = (X_0(t), X_1(t), X_2(t))$ alourdie en "0".

— $G_1(t, \tilde{\mathbf{z}}) = \gamma_{10}(t) + \gamma_{11}(t)z_1$ de loi Bernoulli alourdie en "0"

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0)) = \gamma_{10}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 1, 0)) = \gamma_{11}(t). \end{cases}$$

— $G_2(t, \tilde{\mathbf{z}}) = \gamma'_{10}(t) + \gamma'_{11}(t)z_1 + \gamma'_{12}(t)z_2$ de loi 2-monomiale alourdie en "0"

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0)) = \gamma'_{10}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 1, 0)) = \gamma'_{11}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 1)) = \gamma'_{12}(t). \end{cases}$$

2. **Le cas 2.3 de (4.31)** $\lambda_{12} = \lambda_{21} = 0$ et $\mu_1 = \mu_2$, c.à.d. $A = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_1 \end{pmatrix}$

La décomposition de A en bloc est déjà faite et donne

$$\int_0^t e^{-As} ds = \begin{pmatrix} \frac{(1-e^{-\Lambda_1 t})}{\Lambda_1} & 0 \\ 0 & \frac{(1-e^{-\Lambda_1 t})}{\Lambda_1} \end{pmatrix}, \nu = - \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathcal{I}(0) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

c.à.d.

$$\begin{aligned} \mathcal{I}(t) &= e^{\Lambda_1 t} \begin{pmatrix} y_1 + \frac{e^{-\Lambda_1 t} - 1}{\Lambda_1} \\ y_2 + \frac{e^{-\Lambda_1 t} - 1}{\Lambda_1} \end{pmatrix} \\ &= \begin{pmatrix} e^{\Lambda_1 t} & 0 \\ 0 & e^{\Lambda_1 t} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{\Lambda_1}(1 - e^{\Lambda_1 t}) \\ \frac{1}{\Lambda_1}(1 - e^{\Lambda_1 t}) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \end{aligned}$$

autrement dit

$$\begin{aligned} \begin{pmatrix} G_1(t, \tilde{\mathbf{z}}) \\ G_2(t, \tilde{\mathbf{z}}) \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} &= \begin{pmatrix} e^{-\Lambda_1 t} & 0 \\ 0 & e^{-\Lambda_1 t} \end{pmatrix} \begin{pmatrix} z_1 - \frac{1}{\Lambda_1}(1 - e^{\Lambda_1 t}) \\ z_2 - \frac{1}{\Lambda_1}(1 - e^{\Lambda_1 t}) \end{pmatrix} \\ &= \begin{pmatrix} e^{-\Lambda_1 t} z_1 - \frac{1}{\Lambda_1}(e^{-\Lambda_1 t} - 1) \\ e^{-\Lambda_1 t} z_2 - \frac{1}{\Lambda_1}(e^{-\Lambda_1 t} - 1) \end{pmatrix}. \end{aligned}$$

Conclusion :

Pour tout $i \geq 1$, $G_i(t, \tilde{\mathbf{z}})$ est une loi de Bernoulli de $X_i(t)$ alourdie en “0”.

(a) $G_1(t, \tilde{\mathbf{z}}) = \delta_{10}(t) + \delta_{11}(t)z_1$, correspond à la loi

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0)) = \delta_{10}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 1, 0)) = \delta_{11}(t). \end{cases}$$

(b) $G_2(t, \tilde{\mathbf{z}}) = \delta'_{10}(t) + \delta'_{12}(t)z_2$, correspond à la loi

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0)) = \delta'_{10}(t), \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 1)) = \delta'_{12}(t). \end{cases}$$

4.5.1 Simulations de l'exemple de branchement 1 mère-2 filles

Ici, on simule un branchement avec $K = 3$ populations. Dans ce modèle, la première population (notée de type “0”, en orange dans le graphe 4.8), jouera le rôle “mère” : c’est une population donneuse, c’est-à-dire que, chaque individu de cette population meurt (μ_0) ou donne naissance aux 2 autres avec les taux $\lambda_{01}, \lambda_{02}$, sans en recevoir contrepartie. Elle s’accroît donc de façon autonome par rapport aux autres (λ_{00}). Les deux autres populations (bleu de type “1”, rouge de type “2”), appelées populations “filles” meurent avec des taux de mortalité $(\mu_j)_{j=1,2}$ ou interagissent de façon linéaire entres elles ($\lambda_{12}, \lambda_{21}$).

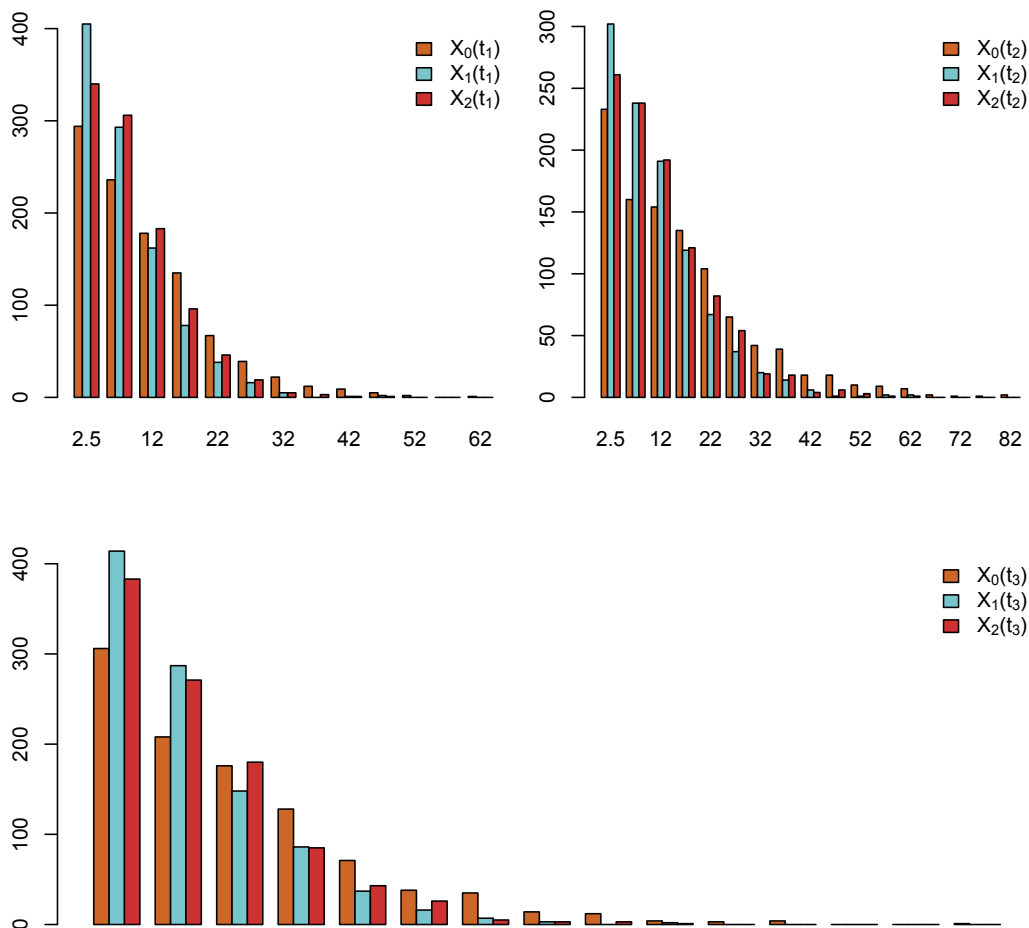


FIGURE 4.8 – Simulations de 3 populations en 3 dates.

La figure 4.8 représente l’histogramme des tailles respectives de 3 populations pour 3 dates de simulation $t = (5, 6, 7)$. La simulation d’une trajectoire comporte 10 000 branchements. On simule 1 000 trajectoires avec les même paramètres et conditions initiales. La taille initiale de chaque population est aléatoire de loi de Poisson de paramètre γ qui suit elle même une loi uniforme sur $[0, 10]$ pour chaque population.

Dans cette simulation, $X_0(0) = 2$, $X_1(0) = 0$, $X_2(0) = 8$. Les paramètres de mort et de branchements suivent des lois uniformes sur $[0, 1]$. Ici, pour la première population : $\mu_0 = 0.13$, $\lambda_{00} = 0.49$, $\lambda_{01} = 0.67$, $\lambda_{02} = 0.52$; tandis que pour les deux autres ($\mu_1 = 0.81$, $\lambda_{12} = 0.05$), ($\mu_2 = 0.42$, $\lambda_{21} = 0.16$).

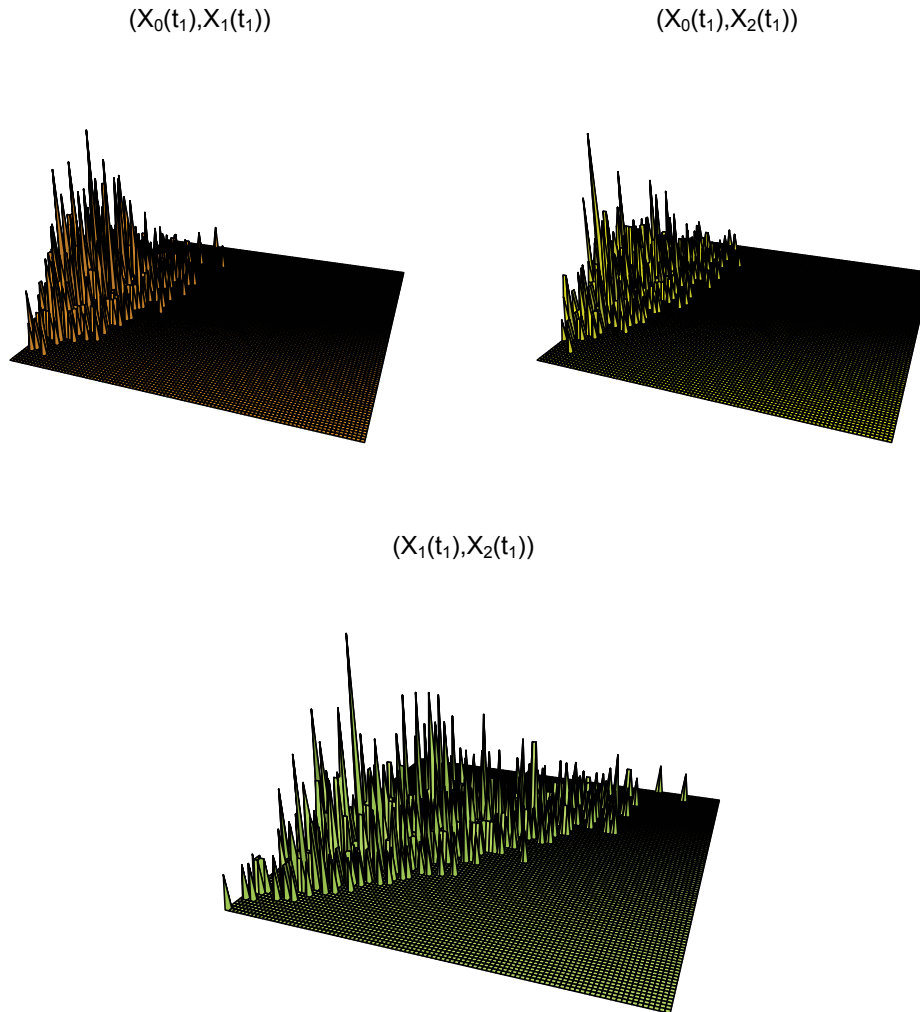


FIGURE 4.9 – Histogrammes 2-D des tailles conjointes de populations 2 à 2.

La figure 4.9 représente les histogrammes 2-D des tailles conjointes de 2 populations $(X_0(t), X_1(t))$, $(X_0(t), X_2(t))$ et $(X_1(t), X_2(t))$ au temps $t = 5$. Ces histogrammes résultent de la simulation de 1 000 trajectoires. Chaque trajectoire comporte 10 000 branchements pour les 3 populations avec les mêmes paramètres et conditions initiales. La taille initiale de chaque population est aléatoire de loi de Poisson de paramètre γ qui suit une loi uniforme sur $[0, 10]$. Dans cette simulation, on a pris $X_0(0) = 8$, $X_1(0) = 0$, $X_2(0) = 3$. Les paramètres de mort et de branchements suivent des lois uniformes sur $[0, 1]$. Ici, pour la première population : $\mu_0 = 0.25$, $\lambda_{00} = 0.29$, $\lambda_{01} = 0.57$, $\lambda_{02} = 0.32$, tandis que pour les deux autres ($\mu_1 = 0.93$, $\lambda_{12} = 0.41$), ($\mu_2 = 0.05$, $\lambda_{21} = 0.90$).

4.6 Sous-modèle du branchement multi-types mère-filles

Nous considérerons un cas particulier plus simple du modèle précédent, où chaque population fournit seulement la population qui suit, et dont l'équation générale est

$$\partial_t G(t, \tilde{\mathbf{z}}) - \sum_{j=0}^K V_j(\tilde{\mathbf{z}}) \partial_{\tilde{\mathbf{z}}_j} G(t, \tilde{\mathbf{z}}) = 0, \quad \tilde{\mathbf{z}} = (z_0, z_1, \dots, z_K) = (z_0, \mathbf{z}),$$

à $K + 1$ populations. La première population est notée de type "0" (population des *mères*). Cette classe s'accroît de façon autonome par rapport aux autres mais fournit "linéairement" sa classe fille "1", c'est à dire

$$V_0(\tilde{\mathbf{z}}) = \mu_0 - \Lambda_0 z_0 + \lambda_{00} z_0^2 + \lambda_{01} z_1$$

Cette dernière fournit linéairement sa classe fille "2". Chaque population $j \geq 2$ fournit linéairement sa population fille $j + 1$ qui suit, autrement dit,

$$\begin{cases} V_1(\tilde{\mathbf{z}}) = \mu_1 - \Lambda_1 z_1 + \lambda_{1,2} z_2 \\ V_2(\tilde{\mathbf{z}}) = \mu_2 - \Lambda_2 z_2 + \lambda_{2,3} z_3 \\ \vdots \\ V_K(\tilde{\mathbf{z}}) = \mu_K - \Lambda_K z_K \end{cases}$$

Les composantes $\mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1(t, \mathbf{y}), \dots, \mathcal{I}_K(t, \mathbf{y}))$ satisfont l'équation différentielle linéaire non homogène :

$$\frac{d\mathcal{I}}{dt}(t, \mathbf{y}) = -\mathbf{V}(\mathcal{I}(t, \mathbf{y})).$$

autrement dit

$$\frac{d\mathcal{I}_j}{dt} = \Lambda_j \mathcal{I}_j - \lambda_{j,j+1} \mathcal{I}_{j+1} - \mu_j, \quad \Lambda_j = \mu_j + \lambda_{j,j+1}.$$

D'où $\frac{d\mathcal{I}}{dt} = A\mathcal{I} + \nu$ où $\nu^T = -(\mu_1, \dots, \mu_K)$ avec A est la matrice :

$$A = \begin{pmatrix} \Lambda_1 & -\lambda_{12} & 0 & \cdots & & 0 \\ 0 & \Lambda_2 & -\lambda_{23} & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & 0 \\ & & & \ddots & \ddots & -\lambda_{K-1K} \\ 0 & & \cdots & 0 & \Lambda_K & \end{pmatrix},$$

ce qui donne $\nu = A\mathbb{I}$

La solution générale de cette dernière équation est classique

$$\mathcal{I}(t, \mathbf{y}) = e^{At} \left[\mathbf{y} + \left(\int_0^t e^{-As} ds \right) \nu \right] \quad (4.32)$$

On peut supposer que la matrice A du modèle est inversible. En effet, si ce n'est pas le cas, alors il existerait une ligne nulle : $\exists j \mid \Lambda_j = 0 = \lambda_{jj+1}$. Autrement dit, on considère seulement les composantes jusqu'à la ligne nulle.

Ainsi on trouve :

$$\int_0^t e^{-As} ds = (\mathbb{I} - e^{-At}) A^{-1} \Rightarrow \left(\int_0^t e^{-As} ds \right) \nu = (\mathbb{I} - e^{-At}) A^{-1} \nu$$

autrement dit,

$$\mathcal{I}(t, \mathbf{y}) = e^{At}(\mathbf{y} - \mathbb{I}) + \mathbb{I}.$$

De plus $\mathcal{I}_0(t, \tilde{\mathbf{y}})$ avec $\mathcal{I}_0(0, \tilde{\mathbf{y}}) = y_0$ vérifie l'équation :

$$\mathcal{I}'_0(t, y_0, \tilde{\mathbf{y}}) = -(\mu_0 + \lambda_{01} \mathcal{I}_1(t, \mathbf{y})) - \Lambda_0 \mathcal{I}_0(t, \tilde{\mathbf{y}}) + \lambda_{00} \mathcal{I}_0^2(t, \tilde{\mathbf{y}})$$

Au final, pour un bloc J de type donné, la solution de l'intégrale vectorielle s'écrit

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = e^{Jt} \left[\tilde{\mathcal{I}}(0, \mathbf{y}) + \left(\int_0^t e^{-Js} ds \right) \tilde{\nu} \right] \quad (4.33)$$

Il faut donc chercher les valeurs et les vecteurs propres de la matrice A pour trouver la solution générale. Par ailleurs, on sait que les valeurs propres d'une matrice triangulaire supérieure sont les coefficients de la diagonale, c'est à dire les Λ_j . Il nous reste donc à trouver les vecteurs propres $W^T = (W_1, \dots, W_K)$ associés à Λ_j : $AW = \Lambda_j W$.

On voit que $W_1 = e_1^T = (1, 0, \dots, 0)$, est un vecteur propre. Pour les autres, on opère de la façon suivante :

$$W_2^T = \left(\frac{-\lambda_{12}}{\Lambda_2 - \Lambda_1}, 1, 0, \dots, 0 \right)$$

$$W_3^T = \left(\frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_3 - \Lambda_2)}, \frac{-\lambda_{23}}{\Lambda_3 - \Lambda_2}, 1, 0, \dots, 0 \right)$$

$$W_4^T = \left(\frac{-\lambda_{12}\lambda_{23}\lambda_{34}}{(\Lambda_4 - \Lambda_1)(\Lambda_4 - \Lambda_2)(\Lambda_4 - \Lambda_3)}, \frac{\lambda_{23}\lambda_{34}}{(\Lambda_4 - \Lambda_2)(\Lambda_4 - \Lambda_3)}, \frac{-\lambda_{34}}{\Lambda_4 - \Lambda_3}, 1, 0, \dots, 0 \right)$$

et le $K^{\text{ème}}$ vecteur propre s'écrit sous la forme suivante :

$$W_K^T = \left((-1)^{K+1} \prod_{i=1}^{K-1} \frac{\lambda_{ii+1}}{(\Lambda_K - \Lambda_i)}, (-1)^{K+2} \prod_{i=2}^{K-1} \frac{\lambda_{ii+1}}{(\Lambda_K - \Lambda_i)}, \dots, \frac{-\lambda_{K-1K}}{(\Lambda_K - \Lambda_{K-1})}, 1 \right)$$

d'où la matrice de passage P de A :

$$P = \begin{pmatrix} 1 & \frac{-\lambda_{12}}{\Lambda_2 - \Lambda_1} & \frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_3 - \Lambda_2)} & \cdots & (-1)^{K+1} \prod_{i=1}^{K-1} \frac{\lambda_{ii+1}}{(\Lambda_K - \Lambda_i)} \\ 0 & 1 & \frac{-\lambda_{23}}{\Lambda_3 - \Lambda_2} & \ddots & (-1)^{K+2} \prod_{i=2}^{K-1} \frac{\lambda_{ii+1}}{(\Lambda_K - \Lambda_i)} \\ & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ & & & \ddots & 1 & \frac{-\lambda_{K-1K}}{(\Lambda_K - \Lambda_{K-1})} \\ 0 & & \cdots & & 0 & 1 \end{pmatrix}$$

Donc on a la formule des coefficients de la matrice de passage P associée a la matrice triangulaire supérieure A ,

1. $P_{ij} = (-1)^{j+i} \prod_{l=i}^{j-1} \frac{\lambda_{l+1}}{(\Lambda_j - \Lambda_l)}$ si $i = j$;
2. $P_{ii} = 1$ si $i = j$;
3. $P_{ij} = 0$ si $i > j$.

Comme P est une matrice triangulaire supérieure avec des coefficients diagonaux tous égaux à 1, la $(1, K)$ coordonnée de la matrice inverse P^{-1} vaut

$$P_{1K}^{-1} = (-1)^{1+K} \left[\Pi_{SD}(P) + \sum_{i=1}^{K-2} (-1)^i \sum_{2 \leq j_1 < \cdots < j_i \leq K-1} \Pi_{SD}(\text{Dec}(j_1, \dots, j_i)(P)) \right],$$

où nous avons défini $\Pi_{SD}(P)$ comme le produit des coefficients de la deuxième diagonale supérieure de P

$$\Pi_{SD}(P) = \prod_{i=1}^{n-1} P_{i,i+1},$$

ainsi que la matrice $\text{Dec}(j_1, \dots, j_i)(P)$, comme la sous-matrice de P dans laquelle nous avons supprimé les lignes et colonnes j_1, \dots, j_i . Pour déterminer les P_{iK}^{-1} , pour $i > 1$, les coefficients de la matrice inverse P^{-1} vérifient :

1. $P_{ii}^{-1} = 1 \quad \forall i$
2. $P_{ij}^{-1} = 0$ si $i > j$;
3. $P_{ij}^{-1} = (-1)^{i+j} \frac{\prod_{l=i}^{j-1} \lambda_{l+1}}{\prod_{l=i+1}^j (\Lambda_l - \Lambda_i)}$ si $i < j$.

Donc la matrice P^{-1} s'écrit sous la forme suivante

$$P^{-1} = \begin{pmatrix} 1 & \frac{\lambda_{12}}{\Lambda_2 - \Lambda_1} & \frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} & \cdots & \cdots & \frac{\prod_{l=1}^{K-1} \lambda_{l+1}}{\prod_{l=2}^j (\Lambda_l - \Lambda_1)} \\ 0 & 1 & \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} & \ddots & \cdots & \prod_{i=2}^{K-1} \frac{\prod_{l=2}^{K-1} \lambda_{l+1}}{\prod_{l=3}^j (\Lambda_l - \Lambda_2)} \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 & \frac{\lambda_{K-1K}}{(\Lambda_K - \Lambda_{K-1})} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Comme il y a K vecteurs propres distinctes, la matrice de Jordan J est diagonale, ses coefficients sont les valeurs propres Λ_j . Son exponentielle est simple, et s'écrit sous la forme suivante

$$e^{-Js} = \begin{pmatrix} e^{-\Lambda_1 s} & 0 & \cdots & 0 \\ 0 & e^{-\Lambda_2 s} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & e^{-\Lambda_K s} \end{pmatrix}.$$

On fait le changement de variables pour $\tilde{\nu}$

$$\tilde{\nu} = P^{-1}\nu = \begin{pmatrix} \mu_1 + \frac{\lambda_{12}\mu_2}{\Lambda_2 - \Lambda_1} + \frac{\lambda_{12}\lambda_{23}\mu_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} + \cdots + \mu_K \frac{\prod_{l=1}^{K-1} \lambda_{l+1}}{\prod_{l=2}^j (\Lambda_l - \Lambda_1)} \\ \mu_2 + \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} + \cdots + \mu_K \frac{\prod_{l=2}^{K-1} \lambda_{l+1}}{\prod_{l=3}^j (\Lambda_l - \Lambda_2)} \\ \vdots \\ \mu_K \end{pmatrix},$$

ce qui nous donne

$$\left(\int_0^t e^{-Js} ds \right) \tilde{\nu} = \begin{pmatrix} \left(\frac{1-e^{-\Lambda_1 t}}{\Lambda_1} \right) \left(\mu_1 + \frac{\lambda_{12}\mu_2}{\Lambda_2 - \Lambda_1} + \frac{\lambda_{12}\lambda_{23}\mu_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} + \cdots + \mu_K \frac{\prod_{l=1}^{K-1} \lambda_{l+1}}{\prod_{l=2}^j (\Lambda_l - \Lambda_1)} \right) \\ \left(\frac{1-e^{-\Lambda_2 t}}{\Lambda_2} \right) \left(\mu_2 + \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} + \cdots + \mu_K \frac{\prod_{l=2}^{K-1} \lambda_{l+1}}{\prod_{l=3}^j (\Lambda_l - \Lambda_2)} \right) \\ \vdots \\ \mu_K \left(\frac{1-e^{-\Lambda_K t}}{\Lambda_K} \right) \end{pmatrix} \quad (4.34)$$

et

$$\tilde{\mathcal{I}}(0, \mathbf{y}) = P^{-1}\mathcal{I}(0, \mathbf{y}) = \begin{pmatrix} y_1 + \frac{\lambda_{12}y_2}{\Lambda_2 - \Lambda_1} + \frac{\lambda_{12}\lambda_{23}y_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} + \cdots + y_K \frac{\prod_{l=1}^{K-1} \lambda_{ll+1}}{\prod_{l=2}^j (\Lambda_l - \Lambda_1)} \\ y_2 + \frac{\lambda_{23}y_3}{\Lambda_3 - \Lambda_2} + \cdots + y_K \frac{\prod_{l=2}^{K-1} \lambda_{ll+1}}{\prod_{l=3}^j (\Lambda_l - \Lambda_2)} \\ \vdots \\ y_K \end{pmatrix} \quad (4.35)$$

en remplaçant (4.34) et (4.35) dans (4.33), on obtient :

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} e^{\Lambda_1 t} y_1 + \frac{(1-e^{\Lambda_1 t})\mu_1}{\Lambda_1} + \frac{\lambda_{12}}{\Lambda_2 - \Lambda_1} \left(\frac{(1-e^{\Lambda_1 t})\mu_2}{\Lambda_1} + y_2 \right) + \cdots + \frac{\prod_{l=1}^{K-1} \lambda_{ll+1}}{\prod_{l=2}^j (\Lambda_l - \Lambda_1)} \left(\frac{(1-e^{\Lambda_1 t})\mu_K}{\Lambda_1} + y_K \right) \\ e^{\Lambda_2 t} y_2 + \frac{(1-e^{\Lambda_2 t})\mu_2}{\Lambda_2} + \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} \left(\frac{(1-e^{\Lambda_2 t})\mu_3}{\Lambda_2} + y_3 \right) + \cdots + \frac{\prod_{l=2}^{K-1} \lambda_{ll+1}}{\prod_{l=3}^j (\Lambda_l - \Lambda_2)} \left(\frac{(1-e^{\Lambda_2 t})\mu_K}{\Lambda_2} + y_K \right) \\ \vdots \\ e^{\Lambda_K t} y_K + \frac{(1-e^{\Lambda_K t})\mu_K}{\Lambda_K} \end{pmatrix}$$

Pour avoir la solution générale $\mathcal{I}(t, \mathbf{y})$ (formule (4.32)), il suffit de faire le changement de variable

$$\mathcal{I}(t, \mathbf{y}) = P\tilde{\mathcal{I}}(t, \mathbf{y}).$$

4.7 Un exemple du sous-modèle : branchement 4-types

Nous considérons $K = 3$ populations du modèle :

$$V_0(z_0, z_1) = \mu_0 - \Lambda_0 z_0 + \lambda_{0,0} z_0^2 + \lambda_{0,1} z_1. \quad (4.36)$$

et

$$\begin{cases} V_1(z_1, z_2) = \mu_1 - \Lambda_1 z_1 + \lambda_{12} z_2 \\ V_2(z_2, z_3) = \mu_2 - \Lambda_2 z_2 + \lambda_{23} z_3 \\ V_3(z_3) = \mu_3 - \Lambda_3 z_3 \end{cases}, \quad \mathbf{z} = (z_1, z_2, z_3), \quad \Lambda_i = \mu_i + \lambda_{i,i+1}$$

Ainsi $\mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1(t, \mathbf{y}), \mathcal{I}_2(t, \mathbf{y}), \mathcal{I}_3(t, \mathbf{y}))$ satisfait :

$$\frac{d\mathcal{I}}{dt} = A\mathcal{I} + \nu \text{ avec } \mathcal{I}^T(0, \mathbf{y}) = \mathbf{y}, \quad \nu^T = -(\mu_1, \mu_2, \mu_3).$$

Sa solution est donc :

$$\mathcal{I}(t, \mathbf{y}) = e^{At} \left(\mathbf{y} + \left(\int_0^t e^{-As} ds \right) \nu \right)$$

avec la matrice :

$$A = \begin{pmatrix} \Lambda_1 & -\lambda_{12} & 0 \\ 0 & \Lambda_2 & -\lambda_{23} \\ 0 & 0 & \Lambda_3 \end{pmatrix}.$$

Une matrice de passage P (ou P^{-1}) de Jordan pour A est facile à établir. Les valeurs propres de la matrice triangulaire supérieure A sont les coefficients Λ_j , $j \in [1, 3]$. Supposons que les Λ_j sont tous distincts et non nuls.

Selon notre schéma précédent, on prend :

$$P = \begin{pmatrix} 1 & \frac{-\lambda_{12}}{\Lambda_2 - \Lambda_1} & \frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_3 - \Lambda_2)} \\ 0 & 1 & \frac{-\lambda_{23}}{\Lambda_3 - \Lambda_2} \\ 0 & 0 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & \frac{\lambda_{12}}{\Lambda_2 - \Lambda_1} & \frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} \\ 0 & 1 & \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{et } J = \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{pmatrix}.$$

Après les changements de variables on obtient :

$$\begin{aligned} \tilde{\nu} = P^{-1}\nu &= \begin{pmatrix} -\mu_1 - \frac{\lambda_{12}\mu_2}{\Lambda_2 - \Lambda_1} - \frac{\lambda_{12}\lambda_{23}\mu_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} \\ -\mu_2 - \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} \\ -\mu_3 \end{pmatrix}, \\ \tilde{\mathcal{I}}(0, \mathbf{y}) = P^{-1}\mathcal{I}(0, \mathbf{y}) &= \begin{pmatrix} y_1 + \frac{\lambda_{12}y_2}{\Lambda_2 - \Lambda_1} + \frac{\lambda_{12}\lambda_{23}y_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} \\ y_2 + \frac{\lambda_{23}y_3}{\Lambda_3 - \Lambda_2} \\ y_3 \end{pmatrix} \end{aligned} \quad (4.37)$$

et

$$\left(\int_0^t e^{-Js} ds \right) \tilde{\nu} = \begin{pmatrix} \left(\frac{1 - e^{-\Lambda_1 t}}{\Lambda_1} \right) \left(-\mu_1 - \frac{\lambda_{12}\mu_2}{\Lambda_2 - \Lambda_1} - \frac{\lambda_{12}\lambda_{23}\mu_3}{(\Lambda_3 - \Lambda_1)(\Lambda_2 - \Lambda_1)} \right) \\ \left(\frac{1 - e^{-\Lambda_2 t}}{\Lambda_2} \right) \left(-\mu_2 - \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} \right) \\ -\mu_3 \left(\frac{1 - e^{-\Lambda_3 t}}{\Lambda_3} \right) \end{pmatrix}. \quad (4.38)$$

En remplaçant (4.37) et (4.38) dans (4.33), on obtient :

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} e^{\Lambda_1 t} \left(y_1 - \frac{\mu_1}{\Lambda_1} + \frac{\lambda_{12}}{\Lambda_2 - \Lambda_1} \left(y_2 - \frac{\mu_2}{\Lambda_2} \right) + \frac{\lambda_{12}\lambda_{23}}{(\Lambda_2 - \Lambda_1)(\Lambda_3 - \Lambda_1)} \left(y_3 - \frac{\mu_3}{\Lambda_3} \right) \right) + \frac{\mu_1}{\Lambda_1} + \frac{\lambda_{12}}{\Lambda_1(\Lambda_2 - \Lambda_1)} \left(\frac{\mu_3\lambda_{23}}{\Lambda_3 - \Lambda_1} + \mu_2 \right) \\ e^{\Lambda_2 t} \left(y_2 - \frac{\mu_2}{\Lambda_2} + \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} \left(y_3 - \frac{\mu_3}{\Lambda_3} \right) \right) + \frac{\mu_2}{\Lambda_2} + \frac{\mu_3\lambda_{23}}{\Lambda_2(\Lambda_3 - \Lambda_2)} \\ e^{\Lambda_3 t} \left(y_3 - \frac{\mu_3}{\Lambda_3} \right) + \frac{\mu_3}{\Lambda_3} \end{pmatrix}$$

et pour avoir la solution générale $\mathcal{I}(t, \mathbf{y})$ de la formule (4.32), il suffit de faire le changement de variable $\mathcal{I}(t, \mathbf{y}) = P\tilde{\mathcal{I}}(t, \mathbf{y})$:

La solution de la première EDO, qui est la génératrice correspondant à un processus démarrant avec un individu de la première classe, s'écrit

$$\mathcal{I}_1(t, \mathbf{y}) = \Delta_{10}(t) + \Delta_{11}(t)y_1 + \Delta_{12}(t)y_2 + \Delta_{13}(t)y_3$$

avec

$$\begin{aligned} \Delta_{10}(t) &= \frac{-e^{\Lambda_1 t}}{\Lambda_1} \left(\mu_1 + \frac{\lambda_{12}\mu_2}{\Lambda_2\Lambda_1} + \frac{\lambda_{12}\lambda_{23}\mu_3}{(\Lambda_2 - \Lambda_1)(\Lambda_3 - \Lambda_1)} \right) - \frac{\lambda_{12}}{\Lambda_2(\Lambda_2 - \Lambda_1)} \left(\mu_2 - \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} \right) \\ &\quad \left(1 - e^{\Lambda_2 t} \right) - \frac{\lambda_{12}\lambda_{23}(1 - e^{\Lambda_3 t})}{(\Lambda_3 - \Lambda_2)} (\Lambda_3 - \Lambda_1)\Lambda_3, \\ \Delta_{11}(t) &= e^{\Lambda_1 t}, \\ \Delta_{12}(t) &= \lambda_{12} \left(e^{\Lambda_1 t} - \Lambda_2 t \right), \\ \Delta_{13}(t) &= \frac{\lambda_{12}\lambda_{23}e^{\Lambda_1 t}}{(\Lambda_2 - \Lambda_1)(\Lambda_3 - \Lambda_1)} - \frac{\lambda_{12}\lambda_{23}e^{\Lambda_2 t}}{(\Lambda_2 - \Lambda_1)(\Lambda_3 - \Lambda_2)} + \frac{\lambda_{12}\lambda_{23}}{(\Lambda_3 - \Lambda_1)(\Lambda_3 - \Lambda_2)}. \end{aligned}$$

La fonction génératrice de la loi de la classe "2" s'écrit à partir de la forme suivante :

$$\mathcal{I}_2(t, \mathbf{y}) = \Delta'_{20}(y) + \Delta'_{22}(t)y_2 + \Delta'_{23}(t)y_3$$

avec

$$\begin{cases} \Delta'_{20}(t) = \frac{1}{\Lambda_2} \left(\mu_2 + \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} \right) - \frac{\lambda_{23}\mu_3}{\Lambda_3 - \Lambda_2} (1 - e^{\Lambda_2 t}) + \frac{\lambda_{23}\mu_3}{\Lambda_3(\Lambda_3 - \Lambda_2)} (1 - e^{\Lambda_3 t}) \\ \Delta'_{22}(t) = e^{\Lambda_2 t} \\ \Delta'_{23}(t) = \frac{\lambda_{23}}{\Lambda_3 - \Lambda_2} (e^{\Lambda_2 t} - 1) \end{cases}$$

Pour celle de la classe "3", on utilisera

$$\mathcal{I}_3(t, \mathbf{y}) = \Delta''_{30} + \Delta''_{33}(t)y_3$$

avec $\begin{cases} \Delta''_{30} = 1 \\ \Delta''_{33}(t) = \frac{-\mu_3}{\Lambda_3} (1 - e^{\Lambda_3 t}) \end{cases}$ autrement dit :

$$\mathcal{I}(t, \mathbf{y}) = \begin{pmatrix} \Delta_{11}(t) & \Delta_{12}(t) & \Delta_{13}(t) \\ 0 & \Delta'_{22}(t) & \Delta'_{23}(t) \\ 0 & 0 & \Delta''_{33}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} \Delta_{10}(t) \\ \Delta'_{20}(t) \\ \Delta''_{30}(t) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

D'où, pour $\tilde{\mathbf{z}} = (z_0, z_1, z_2, z_3)$

$$\begin{pmatrix} G_1(t, \tilde{\mathbf{z}}) \\ G_2(t, \tilde{\mathbf{z}}) \\ G_3(t, \tilde{\mathbf{z}}) \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \Delta_{11}(t) & \Delta_{12}(t) & \Delta_{13}(t) \\ 0 & \Delta'_{22}(t) & \Delta'_{23}(t) \\ 0 & 0 & \Delta''_{33}(t) \end{pmatrix}^{-1} \begin{pmatrix} z_1 - \Delta_{10}(t) \\ z_2 - \Delta'_{20}(t) \\ z_3 - \Delta''_{30}(t) \end{pmatrix}$$

donc

$$\begin{pmatrix} G_1(t, \tilde{\mathbf{z}}) \\ G_2(t, \tilde{\mathbf{z}}) \\ G_3(t, \tilde{\mathbf{z}}) \end{pmatrix} = \begin{pmatrix} \frac{1}{\Delta_{11}(t)} & -\frac{\Delta_{12}(t)}{\Delta_{11}(t)\Delta'_{22}(t)} & \frac{\Delta_{12}(t)\Delta'_{23}(t) - \Delta_{13}(t)\Delta'_{22}(t)}{\Delta_{11}(t)\Delta'_{22}(t)\Delta''_{33}(t)} \\ 0 & \frac{1}{\Delta'_{22}(t)} & \frac{-\Delta'_{23}(t)}{\Delta'_{22}(t)\Delta''_{33}(t)} \\ 0 & 0 & \frac{1}{\Delta''_{33}(t)} \end{pmatrix} \begin{pmatrix} z_1 - \Delta_{10}(t) \\ z_2 - \Delta'_{20}(t) \\ z_3 - \Delta''_{30}(t) \end{pmatrix}$$

autrement dit

$$\begin{aligned} G_1(t, \tilde{\mathbf{z}}) &= \frac{z_1}{\Delta_{11}(t)} - \frac{\Delta_{12}(t)z_2}{\Delta_{11}(t)\Delta'_{22}(t)} + \left(\frac{\Delta_{12}(t)\Delta'_{23}(t) - \Delta_{13}(t)\Delta'_{22}(t)}{\Delta_{11}(t)\Delta'_{22}(t)\Delta''_{33}(t)} \right) z_3 \\ &\quad + \frac{\Delta_{12}(t)\Delta'_{20}(t) - \Delta_{10}(t)\Delta'_{22}(t)}{\Delta_{11}(t)\Delta'_{22}(t)} + \frac{\Delta''_{30}(t)(\Delta_{13}(t)\Delta'_{22}(t) - \Delta_{12}(t)\Delta'_{23}(t))}{\Delta_{11}(t)\Delta'_{22}(t)\Delta''_{33}(t)} \\ G_2(t, \tilde{\mathbf{z}}) &= \frac{z_2}{\Delta'_{22}(t)} - \frac{\Delta'_{23}(t)z_3}{\Delta'_{22}(t)\Delta''_{33}(t)} + \frac{\Delta'_{23}\Delta''_{30}(t)}{\Delta'_{22}(t)\Delta''_{33}} - \frac{\Delta'_{20}(t)}{\Delta'_{22}(t)} \\ G_3(t, \tilde{\mathbf{z}}) &= \frac{z_3}{\Delta''_{33}(t)} - \frac{\Delta''_{30}(t)}{\Delta''_{33}(t)} \end{aligned}$$

Conclusion :

Pour tout $i \geq 1$, $G_i(t, \tilde{\mathbf{z}})$ est une loi multinomiale de $\mathbf{X}(t) = (X_0(t), X_1(t), X_2(t), X_3(t))$ de masse totale 1, alourdie en “0” = (0, 0, 0, 0) :

— $G_1(t, \tilde{\mathbf{z}})$ de loi

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 0)) = \frac{\Delta_{12}(t)\Delta'_{20}(t) - \Delta_{10}(t)\Delta'_{22}(t)}{\Delta_{11}(t)\Delta'_{22}(t)} + \frac{\Delta''_{30}(t)(\Delta_{13}(t)\Delta'_{22}(t) - \Delta_{12}(t)\Delta'_{23}(t))}{\Delta_{11}(t)\Delta'_{22}(t)\Delta''_{33}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 1, 0, 0)) = \frac{1}{\Delta_{11}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 1, 0)) = -\frac{\Delta_{12}(t)}{\Delta_{11}(t)\Delta'_{22}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 1)) = \frac{\Delta_{12}(t)\Delta'_{23}(t) - \Delta_{13}(t)\Delta'_{22}(t)}{\Delta_{11}(t)\Delta'_{22}(t)\Delta''_{33}(t)}. \end{cases}$$

— $G_2(t, \tilde{\mathbf{z}})$ de loi

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 0)) = \frac{\Delta'_{23}\Delta''_{30}(t)}{\Delta'_{22}(t)\Delta''_{33}} - \frac{\Delta'_{20}(t)}{\Delta'_{22}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 1, 0)) = \frac{1}{\Delta'_{22}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 1)) = -\frac{\Delta'_{23}(t)}{\Delta''_{33}(t)\Delta'_{22}(t)}. \end{cases}$$

— $G_3(t, \tilde{\mathbf{z}})$ de loi

$$\begin{cases} \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 0)) = -\frac{\Delta''_{30}(t)}{\Delta''_{33}(t)}, \\ \mathbb{P}(\mathbf{X}(t) = (0, 0, 0, 1)) = \frac{1}{\Delta''_{33}(t)}. \end{cases}$$

Enfin, il nous reste à chercher la solution pour la classe “0” de (4.36) en partant de :

$$\frac{d\mathcal{I}_0(t, \tilde{\mathbf{y}})}{dt} = \nu(t) - \Lambda_0 \mathcal{I}_0(t, \tilde{\mathbf{y}}) + \lambda_{00} \mathcal{I}_0^2(t, \tilde{\mathbf{y}}) \text{ avec } \nu(t) = \lambda_{01} \mathcal{I}_1(t, \mathbf{y}) + \mu_0.$$

Pour faciliter les calculs, on peut écrire $\mathcal{I}_1(t, \mathbf{y})$ sous la forme suivante :

$$\mathcal{I}_1(t, \mathbf{y}) = \sum_{j=1}^3 \Delta_{1j}(\mathbf{y}) e^{\Lambda_j t} + \Delta_{10}.$$

Donc on aura

$$\nu(t) = \sum_{j=1}^3 a_j(\mathbf{y}) e^{\Lambda_j t} + b \quad \text{avec} \quad \begin{cases} a_j(\mathbf{y}) = \lambda_{01} \Delta_{1j}(\mathbf{y}) \\ b = \lambda_{01} \Delta_{10} + \mu_0 \end{cases}.$$

4.7.1 Simulations de l'exemple du sous-modèle : branchement 4-types

Dans ce modèle ($K = 4$ populations : 1 mère-3 filles), la première (notée de type "0", en rose dans le graphe 4.10), jouera le rôle de "mère". C'est la population donneuse, c'est-à-dire que chaque individu de cette population meurt (taux μ_0) se dédouble avec un taux λ_{00} ou alimente seulement la première population avec un taux λ_{01} mais ne reçoit d'aucune autre classe. Elle s'accroît donc de façon autonome par rapport aux autres. Les individus des autres populations "filles" : type "1" (bleu), type "2" (jaune) et type "3" (vert), meurent avec des taux $(\mu_j)_{j=1,2,3}$ ou interagissent de façon hiérarchique, la 1^{ère} population donne à la 2^{ème} avec le taux λ_{12} , et cette dernière donne à la 3^{ème} avec le taux λ_{23} et la 3^{ème} ne donne à aucune autre.

Histogramme 1-D

La figure 4.10 représente les histogrammes des tailles respectives de 4 populations pour 4 dates de simulation $t = (2, 3, 5, 7)$. La simulation d'une trajectoire comporte 10 000 branchements. On simule 1 000 trajectoires avec les même paramètre et conditions initiales. La taille initiale de chaque population est aléatoire de loi de Poisson de paramètre γ qui suit une loi uniforme sur $[0, 10]$ pour chaque population. Les paramètres de mort et de branchement suivent des lois uniformes sur $[0, 1]$. Dans cette simulation, $X_0(0) = 2$, $X_1(0) = 0$, $X_2(0) = 8$, $X_3(0) = 8$. Les paramètres sont : pour la première population : $\mu_0 = 0.13$, $\lambda_{00} = 0.49$, $\lambda_{01} = 0.67$, et pour les autres ($\mu_1 = 0.81$, $\lambda_{12} = 0.05$), ($\mu_2 = 0.42$, $\lambda_{23} = 0.16$) et $\mu_3 = 0.81$.

Histogramme 2-D

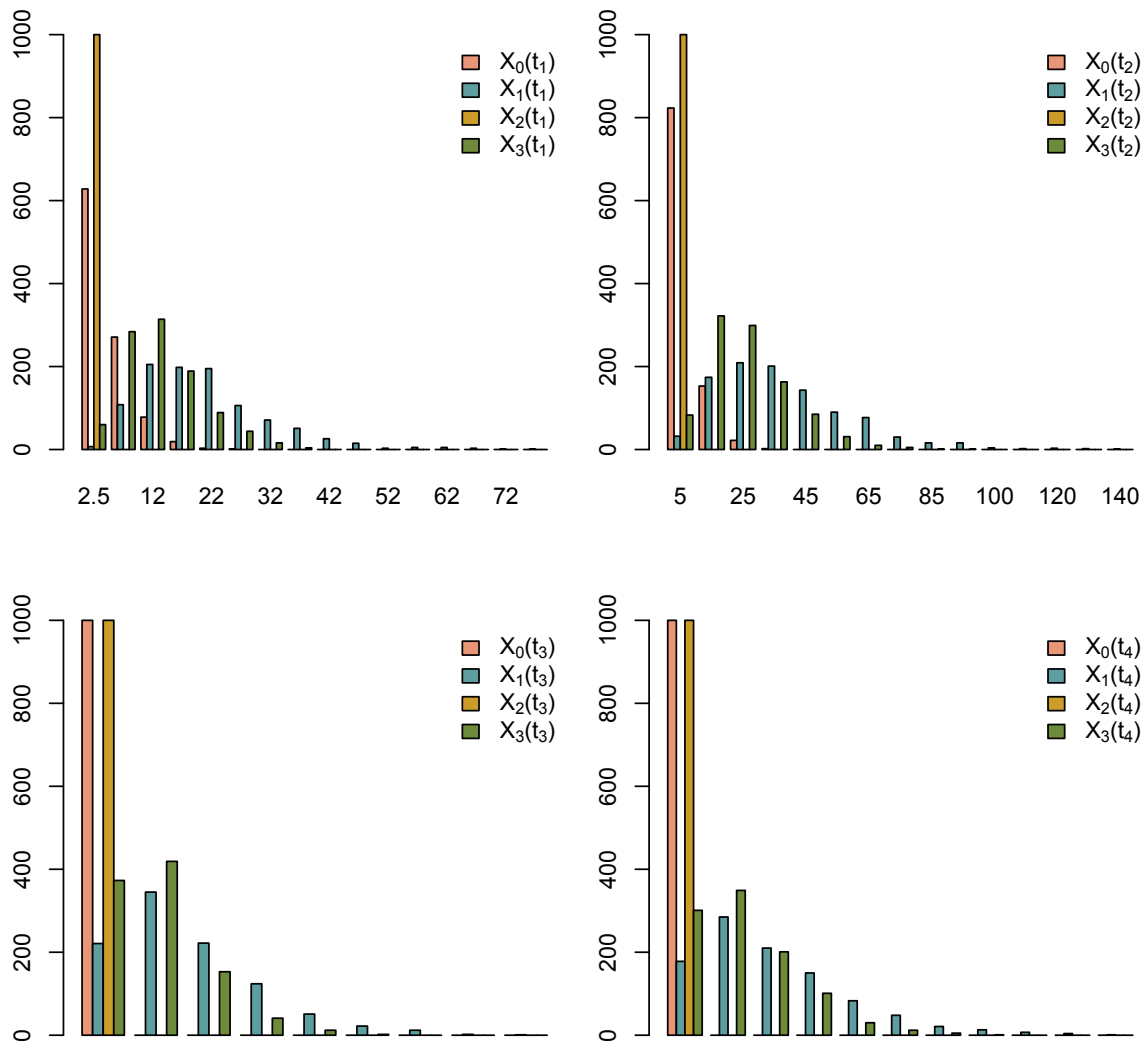


FIGURE 4.10 – Simulations de 4 populations en 4 dates.

Les figures 4.11 et 4.12 représentent respectivement les histogrammes 2-D des tailles conjointes de 2 populations $((X_0(t), X_1(t)), (X_0(t), X_3(t)), (X_1(t), X_3(t))$ et $(X_0(t), X_2(t)), (X_1(t), X_2(t)), (X_2(t), X_3(t)))$ au temps $t = 2$. La taille initiale de chaque population est aléatoire de loi de Poisson de paramètre γ qui suit une loi uniforme sur $[0, 10]$. Dans cette simulation, $X_0(0) = 6$, $X_1(0) = 3$, $X_2(0) = 9$, $X_3(0) = 12$. Les paramètres de décès et de branchements suivent des lois uniformes sur $[0, 1]$. Ici, pour la première population : $\mu_0 = 0.51$, $\lambda_{00} = 0.66$, $\lambda_{01} = 0.70$, et pour les autres ($\mu_1 = 0.01$, $\lambda_{12} = 0.52$), ($\mu_2 = 0.01$, $\lambda_{21} = 0.52$) et $\mu_3 = 0.016$.

Commentaire :

On remarque qu'avec les paramètres égaux les classes 1 et 2 ne se comportent pas de façon identique. La première population reçoit de la population mère, mais ne restitue qu'une partie à la deuxième population, qui, en raison de sa mortalité principalement,

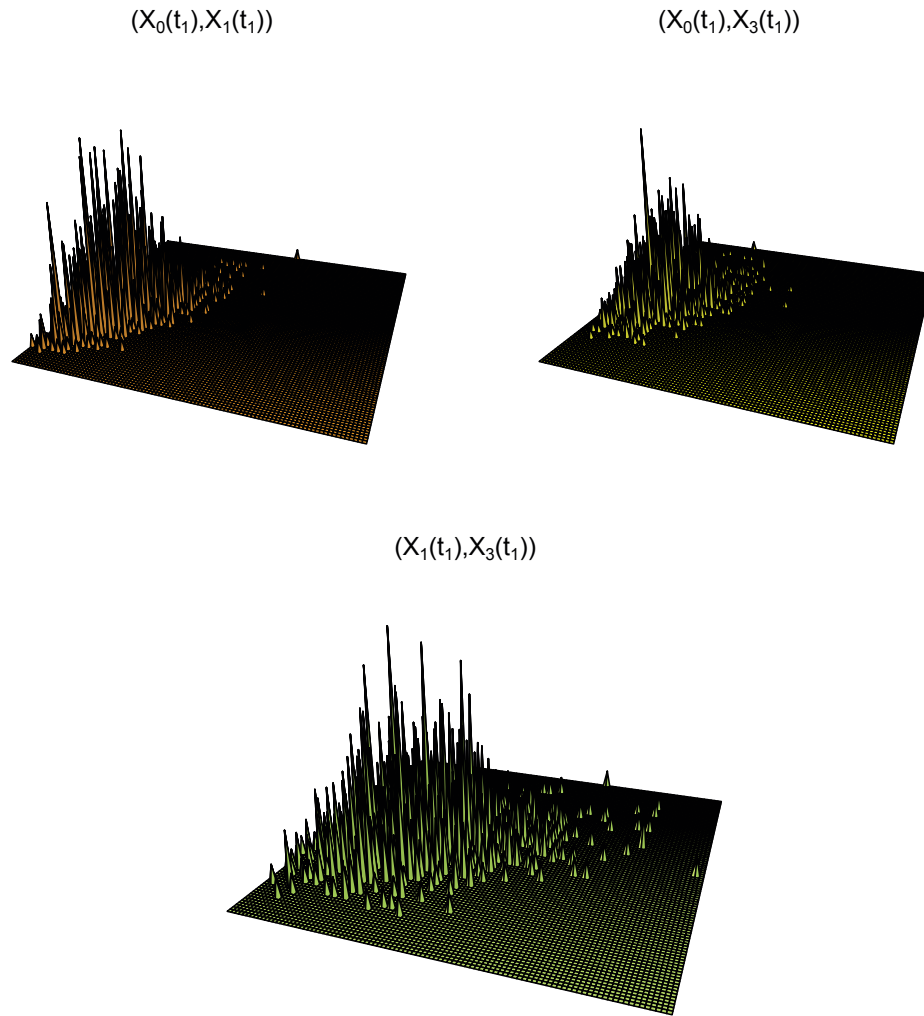


FIGURE 4.11 – Histogrammes 2-D des tailles conjointes de populations $(X_0(t), X_1(t))$, $(X_0(t), X_3(t))$ et $(X_1(t), X_3(t))$.

est d'effectif nul en général. La troisième population (ultime) est assez conséquente car son taux de mortalité est faible.

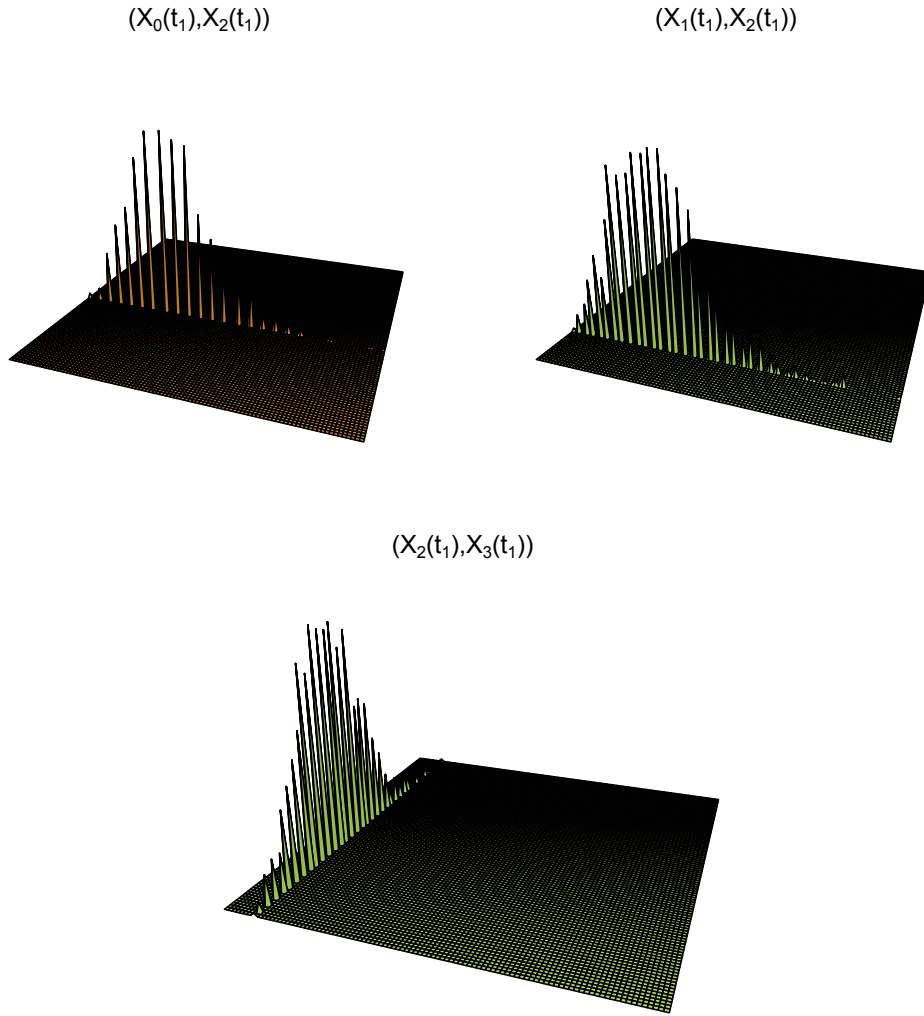


FIGURE 4.12 – Histogrammes 2-D des tailles conjointes de 2 populations $(X_0(t), X_2(t))$, $(X_1(t), X_2(t))$ et $(X_2(t), X_3(t))$.

4.8 Branchement uni-type issu d'un branchement multi-types

Nous allons considérer un cas particulier l'équation générale :

$$\partial_t G(t, \mathbf{z}) + \sum_{j=0}^K V_j(\mathbf{z}) \partial_{z_j} G(t, \mathbf{z}) = 0, \quad \mathbf{z} = (z_0, z_1, \dots, z_K)$$

. avec $K = 3$ populations. La première population de type "0" , vérifie

$$V_0(\mathbf{z}) = \mu_0 - (\mu_0 + \lambda_0)z_0 + \lambda_0 z_0^2.$$

Cette classe, évolue de façon autonome avec une dynamique simple de reproduction quadratique, comme cela a été vu au chapitre 3 (uni-type). D'un autre côté, les deux autres populations fournissent la population "0" linéairement mais sans se reproduire, en d'autres termes cela correspond à

$$\begin{aligned} V_1(\mathbf{z}) &= \mu_1 - \Lambda_1 z_1 + \lambda_{12} z_2 + \lambda_{10} z_0 \\ V_2(\mathbf{z}) &= \mu_2 - \Lambda_2 z_2 + \lambda_{21} z_1 + \lambda_{20} z_0 \end{aligned}$$

avec $\mathbf{z} = (z_1, z_2)$, $\Lambda_i = \mu_i + \sum_{j \neq i}^2 \lambda_{i,j}$.

Utlisant les résultats du chapitre 3, nous obtenons la distribution de probabilité de $X_0(t)$ qui est un processus de branchement 1D quadratique lorsque $X_1(0) = X_2(0) = 0$. Donc si $X_0(0) = n$, on obtient

$$\begin{aligned} - G(t, z) &= \left(\frac{\lambda t - (\lambda t - 1)z}{1 + \lambda t - \lambda t z} \right)^n \quad \text{si } \mu = \lambda, \\ - G(t, z) &= \left(\frac{\mu(e^{\delta t} - 1) - (\mu e^{\delta t} - \lambda)z}{\lambda e^{\delta t} - \mu - (\mu e^{\delta t} - \lambda)z} \right)^n \quad \text{si } \delta = \lambda - \mu \neq 0 \end{aligned}$$

Concernant les autres composantes, on considère, les deux autres intégrales premières $\mathcal{I}^T(t, \mathbf{y}) = (\mathcal{I}_1(t, \mathbf{y}), \mathcal{I}_2(t, \mathbf{y}))$. Elles satisfont l'équation différentielle linéaire non-homogène :

$$\frac{d\mathcal{I}}{dt} = A\mathcal{I} + \nu_0(t) \quad (4.39)$$

avec $\mathcal{I}(0, \mathbf{y}) = \mathbf{y} = (y_1, y_2)$ et $\nu_0^T(t) = \lambda_{10}\mathcal{I}_0(t, y_0) + \mu_1, \lambda_{20}\mathcal{I}_0(t, y_0) + \mu_2$ et avec A , la matrice

$$A = \begin{pmatrix} -\Lambda_1 & \lambda_{12} \\ \lambda_{21} & -\Lambda_2 \end{pmatrix} \quad \Lambda_i = \mu_i + \sum_{j \neq i}^2 \lambda_{i,j}, \quad i = 1, 2.$$

La solution générale de l'équation (4.39) est classique et s'écrit :

$$\mathcal{I}(t, \mathbf{y}) = e^{At} \left(\mathcal{I}(0, \mathbf{y}) + \int_0^t e^{-As} \nu_0(s) ds \right) \quad (4.40)$$

La solution finale implique la réduction de A en blocs de Jordan qui révélera au final la structure des interactions entre les différentes populations. Il s'agit donc de trouver une matrice inversible P , telle que

$$A = P^{-1}JP, \quad \text{ce qui donne } e^{-As} = P^{-1}e^{-Js}P,$$

et de résoudre alors le problème séparément pour chaque bloc de Jordan composant la matrice J .

Les fonctions $\tilde{\mathcal{I}}(t, \mathbf{y}) = P\mathcal{I}(t, \mathbf{y})$ et $\tilde{\nu}_0(t) = P\nu_0(t)$ vérifient

$$\begin{aligned} d\tilde{\mathcal{I}}(t, \mathbf{y}) &= J\tilde{\mathcal{I}}(t, \mathbf{y}) + \tilde{\nu}_0(t) \\ \tilde{\mathcal{I}}(t, \mathbf{y}) &= e^{Jt} \left(\tilde{\mathcal{I}}(0, \mathbf{y}) + \int_0^t e^{-Js} \tilde{\nu}_0(s) ds \right) \end{aligned} \quad (4.41)$$

A chaque valeur propre α de A , on peut associer un ou deux blocs de Jordan. Ainsi, pour les valeurs et les vecteurs propres, on étudiera le déterminant

$$|A - \alpha\mathbf{I}| = \alpha^2 - (\Lambda_1 + \Lambda_2)\alpha + \Lambda_1\Lambda_2 - \lambda_{12}\lambda_{21}$$

et le discriminant $\Delta = (\Lambda_1 + \Lambda_2)^2 + 4\lambda_{12}\lambda_{21}$. Nous considérons alors les deux cas suivants :

Cas 1 : Deux racines distinctes, *i.e.* $\Delta > 0$. Nous avons alors

$$\alpha_{12} = \frac{1}{2}(-\Lambda_1 - \Lambda_2 \pm \sqrt{\Delta})$$

La décomposition de A en blocs de Jordan donne :

$$P = \begin{pmatrix} \frac{\Lambda_2 - \alpha_1}{\lambda_{21}} & \frac{\Lambda_2 - \alpha_2}{\lambda_{21}} \\ 1 & 1 \end{pmatrix}$$

$$P^{-1} = \begin{pmatrix} \frac{\lambda_{21}}{\sqrt{\Delta}} & \frac{-\Lambda_2 - \alpha_2}{\lambda_{21}} \\ -\frac{\lambda_{21}}{\sqrt{\Delta}} & \frac{\Lambda_2 - \alpha_1}{\lambda_{21}} \end{pmatrix}$$

et $J = \begin{pmatrix} \alpha_1 & 0 \\ 1 & \alpha_2 \end{pmatrix}$. Ainsi,

$$\int_0^t e^{-J s} \tilde{\nu}_0(s) ds = \begin{pmatrix} \frac{1}{\alpha_1} ((e^{-\alpha_1 t} - 1)(\lambda_{10} + \mu) + \frac{\lambda_{1,0}}{\lambda_0} e^{\frac{-\alpha_1}{\lambda_0 y'_0}} [E_1\left(\frac{\alpha_1}{\lambda_0 y'_0}\right) - E_1\left(\frac{\alpha_1}{\lambda_0 y'_0 + \alpha_1 t}\right)]) \\ \frac{1}{\alpha_2} (e^{-\alpha_2 t} - 1)(\lambda_{2,0} + \mu) + \frac{\lambda_{2,0}}{\lambda_0} e^{\frac{-\alpha_2}{\lambda_0 y'_0}} [E_1\left(\frac{\alpha_2}{\lambda_0 y'_0}\right) - E_1\left(\frac{\alpha_2}{\lambda_0 y'_0 + \alpha_2 t}\right)] \end{pmatrix}$$

$$\tilde{\mathcal{I}}(0, \mathbf{y}) = P\mathcal{I}(0, \mathbf{y}) = \begin{pmatrix} \frac{\Lambda_2 - \alpha_1}{\lambda_{2,1}} & \frac{\Lambda_2 - \alpha_2}{\lambda_{21}} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\lambda_{21}} [(\Lambda_2 - \alpha_1)y_1 + (\Lambda_2 - \alpha_2)y_2] \\ y_1 + y_2 \end{pmatrix}$$

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} \frac{e^{\alpha_1 t}}{\lambda_{21}} [(\Lambda_2 - \alpha_1)y_1 + (\Lambda_2 - \alpha_2)y_2] + \frac{(1 - e^{\alpha_1 t})(\lambda_{10} + \mu)}{\alpha_1} + \frac{\lambda_{1,0}}{\lambda_0} e^{\alpha_1(t - \frac{1}{\lambda_0 y'_0})} [E_1\left(\frac{\alpha_1}{\lambda_0 y'_0}\right) - E_1\left(\frac{\alpha_1}{\lambda_0 y'_0 + \alpha_1 t}\right)] \\ e^{\alpha_2 t}(y_1 + y_2) + \frac{1}{\alpha_2}(\lambda_{20} + \mu)(1 - e^{\alpha_2 t}) + \frac{\lambda_{2,0}}{\lambda_0} e^{\alpha_2(t - \frac{1}{\lambda_0 y'_0})} [E_1\left(\frac{\alpha_2}{\lambda_0 y'_0}\right) - E_1\left(\frac{\alpha_2}{\lambda_0 y'_0 + \alpha_2 t}\right)] \end{pmatrix}$$

La solution de l'EDO $\mathcal{I}(t, \mathbf{y})$, est donné par $\mathcal{I}(t, \mathbf{y}) = P^{-1}\tilde{\mathcal{I}}(t, \mathbf{y})$

Cas 2 : Valeur propre double , avec 2 blocs de Jordan. Dans ce cas $\Delta = 0$, c.à.d.

Cas 2.1 : $\lambda_{12} = 0$ et $\Lambda_1 = \Lambda_2$, ou **Cas 2.2 :** $\lambda_{21} = 0$ et $\Lambda_1 = \Lambda_2$.

1- Prenons le **Cas 2.1 :** $\lambda_{12} = 0$ et $\Lambda_1 = \Lambda_2$, avec la double valeur propre $\alpha = \Lambda_1$.

La décomposition de A en blocs est déjà faite et donne la matrice de passage

$$P = \begin{pmatrix} 0 & \frac{-1}{\lambda_{21}} \\ 1 & 0 \end{pmatrix}, \text{ d'où}$$

$$P^{-1} = \begin{pmatrix} 0 & 1 \\ -\lambda_{21} & 0 \end{pmatrix}.$$

Soit $\gamma = \lambda_0 y'_0$ avec $y'_0 = 1 - y$. Alors l'intégrale $\int_0^t e^{-J s} \tilde{\nu}_0(s) ds$ vaut :

$$\int_0^t e^{-J s} \tilde{\nu}_0(s) ds = \begin{pmatrix} \frac{1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) (\lambda_{1,0} + \mu_1) + \frac{\lambda_{10}}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] - \frac{t^2}{2} (\lambda_{20} + \mu_2) + \\ \lambda_{20} y'_0 \left[\ln(\gamma + 1) \left(t \left(1 + \frac{1}{\gamma} \right) \right) + \frac{1}{(\gamma)^2} \right] \\ \frac{1}{\Lambda_1} (\lambda_{20} + \mu_2) (e^{-\Lambda_1 t} - 1) + \frac{\lambda_{20}}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \end{pmatrix}.$$

Pour $t = 0$ nous avons

$$\tilde{\mathcal{I}}(0) = P\mathcal{I}(0) = \begin{pmatrix} -\frac{y_2}{\lambda_{21}} \\ y_1 \end{pmatrix}$$

Ainsi

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = e^{Jt} \begin{pmatrix} A \\ B \end{pmatrix} \text{ avec } e^{Jt} = \begin{pmatrix} e^{\Lambda_1 t} & t \\ 0 & e^{\Lambda_1 t} \end{pmatrix},$$

et

$$A = -\frac{y_2}{\lambda_{21}} + \frac{1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) (\lambda_{1,0} + \mu_1) + \frac{\lambda_{10}}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \\ - \frac{t^2}{2} (\lambda_{20} + \mu_2) + \lambda_{2,0} y'_0 \left[\ln(\gamma + 1) \left(t \left(1 + \frac{1}{\gamma} \right) \right) + \frac{1}{(\gamma)^2} \right]$$

et

$$B = y_1 + \frac{1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) (\lambda_{2,0} + \mu_2) + \frac{\lambda_{2,0}}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right]$$

Par conséquent, la solution a la forme :

$$\mathcal{I}(t, \mathbf{y}) = P^{-1} \tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} B e^{\Lambda_1 t} \\ -\lambda_{21} (A e^{\Lambda_1 t} + B t) \end{pmatrix}$$

- 2- Dans le **Cas 2.2** : $\lambda_{12} = \lambda_{21}$ et $\Lambda_1 = \Lambda_2$, la décomposition de A en blocs est déjà faite et donne :

$$\int_0^t e^{-J s} \tilde{\nu}_0(s) ds = \begin{pmatrix} \frac{1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) (\lambda_{1,0} + \mu_1) - \frac{\lambda_{10} y'_0}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \\ \frac{1}{\Lambda_1} (e^{-\Lambda_1 t} - 1) (\lambda_{2,0} + \mu_2) + \frac{\lambda_{20}}{\lambda_0} e^{-\frac{\Lambda_1}{\gamma}} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \end{pmatrix}$$

et

$$\tilde{\mathcal{I}}(t, \mathbf{y}) = \begin{pmatrix} y_1 e^{\Lambda_1 t} + \frac{1}{\Lambda_1} (1 - e^{\Lambda_1 t}) (\lambda_{10} + \mu_1) - \frac{\lambda_{10} y'_0}{\lambda_0} e^{\Lambda_1 (t - \frac{1}{\gamma})} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \\ y_2 e^{\Lambda_1 t} + \frac{1}{\Lambda_1} (1 - e^{\Lambda_1 t}) (\lambda_{20} + \mu_2) + \frac{\lambda_{20}}{\lambda_0} e^{\Lambda_1 (t - \frac{1}{\gamma})} \left[E_1 \left(\frac{\Lambda_1}{\gamma} \right) - E_1 \left(\Lambda_1 t + \frac{\Lambda_1}{\gamma} \right) \right] \end{pmatrix}.$$

Chapitre 5

Conclusion et perspectives

Sommaire

5.1	Approche statistique par Max-Vraisemblance	65
5.2	Vraisemblance pour des observations ponctuelles	66
5.3	Inversion de la génératrice	67

Les processus de branchement uni et multi-type sont des modèles mathématiques toujours très étudiés, de par leur complexité et la richesse mathématique qu'ils génèrent, mais aussi par la pertinence de leurs applications, en biologie notamment. Le présent travail approfondit et unifie partiellement les résultats obtenus dans des travaux antérieurs sur la construction de modèles paramétriques. Les résultats de cette thèse sont destinés à être utilisés pour l'inférence statistique sur ces processus ainsi paramétrés, en vue d'éventuelles applications en biologie.

Ce dernier chapitre ouvre des perspectives et tire les conclusions de ce travail. Nous argumentons la pertinence de notre approche par l'interprétation directe des paramètres, qui sont d'un grand intérêt pour les nombreux domaines d'applications. L'intérêt de développer des méthodes d'inférence statistique propres à ces modèles est de grande actualité. Il existe de nombreuses approches d'estimation paramétrique pour les processus de branchement, mais celles-ci sont pour la plupart basées sur des méthodes de moments et de propriétés de martingales [1].

Dans ce travail, nous contribuons à combler le manque flagrant de lois de probabilité spécifiées et paramétrées modélisant les comptages simultanés des individus de plusieurs populations interactives et évoluant dans le temps. Nous proposons ainsi de revenir à la méthode classique de l'estimation par maximum de vraisemblance des paramètres associés à des processus de branchement uni et multi-types en temps continu, homogènes dans le temps. Notre approche statistique ultérieure consistera à inverser théoriquement ou numériquement la fonction génératrice en utilisant le théorème de Cauchy pour les fonctions analytiques et qui permettra d'inférer ces paramètres.

5.1 Approche statistique par Max-Vraisemblance

Pour l'interprétation du modèle de branchement, la paramétrisation des lois de transition $P_{i,j}(t)$ ou de façon équivalente des génératrices $G_i(t, \mathbf{z})$ se fait naturellement par le biais du générateur infinitésimal $Q_{i,j}$. Le modèle est entièrement déterminé par l'ensemble des paramètres entrant dans la définition des fonctions analytiques définissant la dynamique de ces K populations :

$$V_j(\mathbf{z}) = \partial_t G_j(0, \mathbf{z}) = \sum_h Q_{j,h} \mathbf{z}^h \text{ pour } j = 1, \dots, K,$$

et

$$\begin{cases} \partial_t G(t, \mathbf{z}) = \sum_{j=1}^k V_j(\mathbf{z}) \partial_{z_j} G(t, \mathbf{z}), \\ G(0, \mathbf{z}) = \mathbf{z}. \end{cases}, \text{ avec } G(t, \mathbf{z}) = (G_1(t, \mathbf{z}), \dots, G_K(t, \mathbf{z})). \quad (5.1)$$

- $-Q_{ii}$: K -paramètres pour les lois exponentielles du temps de vie d'un individu de la population i

- $a_{ij} = \frac{Q_{ei}}{-Q_{ii}}$: est la probabilité pour qu'un individu de la population i produise $j = (j_1, \dots, j_K)$ descendants à sa mort.
- Les paramètres a_{ij} définissent les fonctions génératrices $g_i(\mathbf{z}) = \sum_{j \in \mathbb{N}^K} a_{i,j} \mathbf{z}^j$ des lois de probabilité des effectifs des naissances instantanées dans chaque type de population.

En général, il y a une infinité de paramètres, sauf si l'on s'intéresse à des classes de processus de branchement paramétrés spécifiquement par un paramètre d'intérêt $\theta = (\theta_1, \dots, \theta_r)$ de dimension finie r , *i.e.* $V_j(\mathbf{z}) = V_j(\mathbf{z} \text{ et } \theta)$, $Q_{ij} = Q_{ij}(\theta)$.

Le problème d'inférence réside ensuite dans l'inversion des fonctions génératrices après avoir résolu théoriquement ou numériquement les équations dynamiques (5.1) afin d'extraire les probabilités correspondantes.

5.2 Vraisemblance pour des observations ponctuelles

Si les observations de $\mathbf{X}(t)$ sont disponibles aux temps t_1, \dots, t_m , la propriété de Markov dit que la loi de ces observations s'écrit :

$$\mathcal{L}(X(t_1), \dots, X(t_m) | X(0)) = \mathcal{L}(X(t_1) | X(0)) \cdots \mathcal{L}(X(t_m) | X(t_{m-1})).$$

L'hypothèse d'homogénéité nous dit aussi que, pour des observations n_j à valeurs dans \mathbb{N}^K

$$X(0) = n_0, X(t_1) = n_1, \dots, X(t_m) = n_m,$$

la vraisemblance s'écrit :

$$V_\theta = P_{n_0, n_1}(t_1 | \theta) P_{n_1, n_2}(t_2 - t_1 | \theta) \cdots P_{n_{m-1}, n_m}(t_m - t_{m-1} | \theta),$$

où $\theta = (\theta_1, \dots, \theta_r)$ est un vecteur des paramètres spécifiant la forme du générateur infinitésimal.

Dans les modèles réguliers, l'estimateur du maximum de vraisemblance s'obtient par la résolution des r équations :

$$\partial_{\theta_l} \log V = \sum_{i=1}^m \frac{\partial_{\theta_l} P_{n_{i-1}, n_i}(t_i - t_{i-1} | \theta)}{P_{n_{i-1}, n_i}(t_i - t_{i-1} | \theta)} = 0, l = 1, \dots, r.$$

Ceci nous amène à calculer dans un premier temps les fonctions génératrices $g_n(t, \mathbf{z})$ des probabilités de transitions $P_{n, \cdot}(t, \cdot | \theta)$ où $n = (n_1, \dots, n_K)$. On sait que

$$g_n(t, \mathbf{z} | \theta) = G^n(t, \mathbf{z} | \theta) = G_1^{n_1}(t, \mathbf{z} | \theta) \times \cdots \times G_K^{n_K}(t, \mathbf{z} | \theta)$$

permet formellement de retrouver les probabilités par dérivation

$$P_{n, m}(t | \theta) = \frac{1}{m!} \partial_{z_1^{m_1} \dots z_K^{m_K}} G^n(t, \mathbf{0} | \theta).$$

Ainsi, pour $\Delta_i = t_i - t_{i-1}$, on doit calculer $P_{n-1}(\Delta_i, n_i | \theta)$ et

$$\partial_{\theta_i} P_{n_{i-1}}(\Delta_i, n_i | \theta) = \frac{1}{n!} \partial_{\theta_r} \partial_z^{n_i} G_{n_{i-1}, \theta}(\Delta_i, 0)$$

Comme conséquence on voit que l'on a intérêt à prendre des t_i équidistants pour faire un seul et même calcul n'impliquant que des $G_i(\Delta, \mathbf{z})$.

5.3 Inversion de la génératrice

Dans le chapitre 3, on a obtenu des formes simples pour les génératrices. Cependant, tel n'est pas toujours le cas, et il faut alors recourir à des méthodes numériques. Nous commençons par les remarques suivantes :

Soit \mathbf{X} est une v.a. sur \mathbb{N}^K de loi $P = (p_i)$, sa fonction génératrice G est une fonction analytique, sur $D(0, 1)^K$.

$$G(\mathbf{z}) = E(\mathbf{z}^{\mathbf{X}}) = \sum P_{\mathbf{n}} \mathbf{z}^{\mathbf{n}} \quad \text{où } \mathbf{n} = (n_1 \cdots n_K) \in \mathbb{N}^K$$

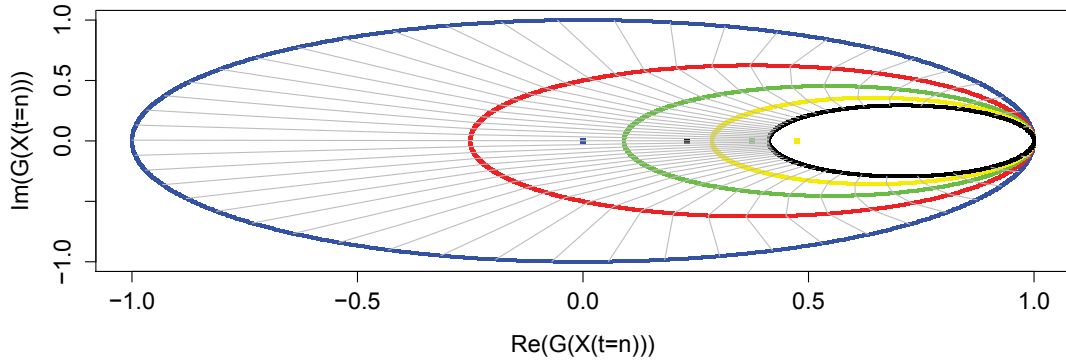


FIGURE 5.1 – Graphe des parties réelles et imaginaires de $G_{X_1}(z)$ (noir), $G_{X_2}(z)$ (jaune), $G_{X_3}(z)$ (vert), $G_{X_4}(z)$ (rouge), $G_{X_5}(z)$ (bleu) pour $z = \frac{1}{2}e^{i\theta}$, $\theta \in [0, 2\pi]$ où Y est le nombre de descendants d'un individu $G_{X_1} = G_Y$.

Nous avons la formule d'inversion multivariée :

$$p_{\mathbf{n}} = \mathbb{P}(\mathbf{X} = \mathbf{n}) = \frac{1}{\mathbf{n}!} \partial_{\mathbf{z}^{\mathbf{n}}} G(\mathbf{0})$$

L'utilisation du théorème de Cauchy pour les fonctions analytiques (uni ou multi-variables) permet précisément de calculer les probabilités correspondantes.

$$P_{n_1 \dots n_K} = \frac{1}{(2\pi i)^K} \int \int \int \frac{G(\xi_1, \dots, \xi_K)}{\xi_1^{n_1+1}, \dots, \xi_K^{n_K+1}} d\xi_1 \dots \xi_K.$$

Pour tout $0 < r_0 < 1$

$$p_{\mathbf{n}} = \frac{1}{(2\pi)^K r_0^{|\mathbf{n}|}} \underbrace{\int_{\gamma_1} \dots \int_{\gamma_K}}_K G(r_0 e^{i\theta_1}, \dots, r_0 e^{i\theta_K}) e^{-i\langle \mathbf{n}, \theta \rangle} d\theta_1 \dots d\theta_K$$

En général, l'intégration peut être prise le long de chemins fermés quelconques $\gamma_i(u)$ contenant 0. Pour simplifier, on a

$$\gamma_i = \text{cercle } C(0, \rho_0), \quad 0 < \rho_0 < 1, \forall i = 1 \dots K.$$

L'approximation sera d'autant meilleure lorsque la dimension K n'est pas très grande et les chemins γ_i simples.

L'intérêt ici est qu'aucune dérivation d'ordre $n = (n_1, \dots, n_K)$ n'est nécessaire, car il s'agit d'intégrales multiples ne faisant intervenir que n . Le choix de r_0 est sûrement important ainsi que celui des pas de discrétisation pour le calcul des intégrales, mais n'ont pas été étudiés ici.

Bibliographie

- [1] S. Asmussen and N. Keiding. Martingale central limit theorems and asymptotic estimation theory for multitype branching processes. *Advances in Applied Probability*, pages 109–129, 1978. [11](#), [65](#)
- [2] K.B. Athreya and P.E. Ney. Branching processes, volume 196 of *die grundlehren der mathematischen wissenschaften*, 1972. [22](#)
- [3] Krishna B Athreya and Peter E Ney. *Branching processes*, volume 196. Springer Science & Business Media, 2012.
- [4] V. Bansaye. Proliferating parasites in dividing cells : Kimmel’s branching model revisited. *The Annals of Applied Probability*, pages 967–996, 2008. [3](#)
- [5] V. Bansaye. Cell contamination and branching processes in a random environment with immigration. *Advances in Applied Probability*, 41(4) :1059–1081, 2009. [3](#)
- [6] V. Bansaye and J. Berestycki. Large deviations for branching processes in random environment. *arXiv preprint arXiv :0810.4991*, 2008. [3](#)
- [7] I.J. Bienaymé. De la loi de multiplication et de la durée des familles. *Soc. Philomat. Paris Extraits, Sér, 5* :37–39, 1845. [4](#)
- [8] N.H. Bingham. Continuous branching processes and spectral positivity. *Stochastic Processes and their Applications*, 4(3) :217–242, 1976. [5](#)
- [9] A. Chen. Ergodicity and stability of generalised markov branching processes with resurrection. *Journal of applied probability*, 39(04) :786–803, 2002. [5](#)
- [10] A. Chen. Uniqueness and extinction properties of generalised markov branching processes. *Journal of mathematical analysis and applications*, 274(2) :482–494, 2002. [5](#)
- [11] A. Chen, J. Li, and N.I. Ramesh. Uniqueness and extinction of weighted markov branching processes. *Methodology and Computing in Applied Probability*, 7(4) :489–516, 2005. [5](#)

-
- [12] A. Chen and E. Renshaw. Markov branching processes regulated by emigration and large immigration. *Stochastic processes and their applications*, 57(2) :339–359, 1995. 7
- [13] K.S. Crump and C.J. Mode. A general age-dependent branching process. i. *Journal of Mathematical Analysis and Applications*, 24(3) :494–508, 1968. 8
- [14] C. Dombry, C. Mazza, and V. Bansaye. Phenotypic diversity and population growth in fluctuating environment : a mbpre approach. *arXiv preprint arXiv :0912.1194*, 2009. 3
- [15] L. Edler. On the first birth and the last death in a generation in a multi-type markov branching process. *Stochastic Processes and their Applications*, 9(2) :175–187, 1979. 9
- [16] A.K. Erlang. Opgave nr. 15. *Mat. Tidsskr. B*, 36, 1929. 4
- [17] R.A. Fisher. *The genetical theory of natural selection : a complete variorum edition*. Oxford University Press, 1999. 4
- [18] F. Galton. Problem 4001. *Educational Times*, 1 :17, 1873. 4
- [19] M. González, R. Martínez, and M. Mota. Multitype population size-dependent branching processes with dependent offspring. *Statistics & probability letters*, 70(2) :145–154, 2004. 6
- [20] A. Grimvall. On the convergence of sequences of branching processes. *The Annals of Probability*, pages 1027–1045, 1974. 4
- [21] P. Guttorp, K. Albertsen, J.F. Steffensen, and E. Kristensen. Three papers on the history of branching processes. *International statistical review*, 63(2) :233–245, 1995. 4
- [22] T.E. Harris. *The theory of branching processes*. Courier Corporation, 2002. 22
- [23] Theodore E Harris. *The theory of branching processes*. Courier Corporation, 2002.
- [24] Ghikman I.I. and Skorohod A.V. *The Theory of Stochastic Processes*. Springer Verlag, 1975.
- [25] C. Jacob. Branching processes : their role in epidemiology. *International journal of environmental research and public health*, 7(3) :1186–1204, 2010. 3
- [26] P. Jagers, F.C. Klebaner, and S. Sagitov. On the path to extinction. *Proceedings of the National Academy of Sciences*, 104(15) :6107–6111, 2007. 6

- [27] P. Jagers, A.N. Lagerås, et al. General branching processes conditioned on extinction are still branching processes. *Electron. Commun. Probab*, 13 :540–547, 2008. 6
- [28] M. Jiřina. Stochastic branching processes with continuous state space. *Czechoslovak Mathematical Journal*, 8(2) :292–313, 1958. 4
- [29] N. Kaplan and T.S. Sudbury, A. and Nilsen. A branching process with disasters. *Journal of Applied Probability*, 12(1) :47–59, 1975. 7
- [30] David G Kendall. Branching processes since 1873. *Journal of the London Mathematical Society*, 1(1) :385–406, 1966.
- [31] B. Klein and P.D.M. Macdonald. The multitype continuous-time markov branching process in a periodic environment. *Advances in Applied Probability*, 12(1) :81–93, 1980. 8
- [32] A.N. Kolmogorov. Zur lösung einer biologischen aufgabe. *Comm. Math. Mech. Chebyshev Univ. Tomsk*, 2(1) :1–12, 1938. 19
- [33] P. Küster. Generalized markov branching processes with state-dependent offspring distributions. *Probability Theory and Related Fields*, 64(4) :475–503, 1983. 9
- [34] J. Lamperti. Continuous state branching processes. *Bulletin of the American Mathematical Society*, 73(3) :382–386, 1967. 4
- [35] J. Lamperti. The limit of a sequence of branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 7(4) :271–288, 1967. 4
- [36] J. Li, A. Chen, and A. G. Pakes. Asymptotic properties of the markov branching process with immigration. *Journal of Theoretical Probability*, 25(1) :122–143, 2012. 7
- [37] Charles J Mode, M Kimmel, and DE Axelrod. Branching processes in biology, 2003.
- [38] C.J. Mode. Multitype age-dependent branching processes and cell cycle analysis. *Mathematical Biosciences*, 10(3-4) :177–190, 1971. 3
- [39] C.J. Mode. *Multitype branching processes : theory and applications*, volume 34. American Elsevier Pub. Co., 1971. 3
- [40] C.J. Mode, M. Kimmel, and D.E. Axelrod. Branching processes in biology, 2003. 22

-
- [41] A. G. Pakes. Extinction and explosion of nonlinear markov branching processes. *Journal of the Australian Mathematical Society*, 82(3) :403–428, 2007. 6
- [42] A. G. Pakes. Critical markov branching process limit theorems allowing infinite variance. *Advances in Applied Probability*, 42(2) :460–488, 2010. 9
- [43] A.G. Pakes. Limit theorems for the simple branching process allowing immigration, i. the case of finite offspring mean. *Advances in Applied Probability*, 11(1) :31–62, 1979. 7
- [44] A.G. Pakes. The markov branching-catastrophe process. *Stochastic processes and their applications*, 23(1) :1–33, 1986. 7
- [45] A.G. Pakes. Ch. 18. biological applications of branching processes. *Handbook of statistics*, 21 :693–773, 2003. 22
- [46] S. Palau and J. C. Pardo. Continuous state branching processes in random environment : The brownian case. *Stochastic Processes and their Applications*, 127(3) :957–994, 2017. 9
- [47] T. A. Ryan. *On age-dependent branching processes*. Cornell University, Sept., 1968. 8
- [48] S. Sagitov. Tail generating functions for extendable branching processes. *Stochastic Processes and their Applications*, 127(5) :1649–1675, 2017. 10
- [49] S. Sagitov and A. Lindo. A special family of galton-watson processes with explosions. In *Branching Processes and Their Applications*, pages 237–254. Springer, 2015. 6
- [50] S. Sagitov and M.C. Serra. Multitype bienaymé–galton–watson processes escaping extinction. *Advances in Applied Probability*, 41(1) :225–246, 2009. 8
- [51] B.A. Sevast’yanov. The theory of branching random processes. *Uspekhi Matematicheskikh Nauk*, 6(6) :47–99, 1951. 22
- [52] Boris Alexandrovich Sevast’yanov. The theory of branching random processes. *Uspekhi Matematicheskikh Nauk*, 6(6) :47–99, 1951.
- [53] B.A. Sevast’yanov. Limit theorems for branching stochastic processes of special form. *Theory of Probability & Its Applications*, 2(3) :321–331, 1957. 6
- [54] M. Slavtchova-Bojkova, P. Trayanov, and S. Dimitrov. Branching processes in continuous time as models of mutations : computational approaches and algorithms. *Computational Statistics & Data Analysis*, 2017. 8

- [55] W.L. Smith and W.E. Wilkinson. On branching processes in random environments. *The Annals of Mathematical Statistics*, pages 814–827, 1969. 9
- [56] Ethier S.N and Kurtz T.G. *Markov Processes. Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.
- [57] D. Tanny. Limit theorems for branching processes in a random environment. *The Annals of Probability*, pages 100–116, 1977. 9
- [58] Howard M Taylor and Samuel Karlin. *An introduction to stochastic modeling*. Academic press, 2014.
- [59] B. Thilaka, B K. Kumar, and A. Vijayakumar. Multitype branching processes with disasters i : The number of particles in the system. *Mathematical and computer modelling*, 28(11) :87–102, 1998. 7
- [60] V. A. Vatutin. A conditional limit theorem for a critical branching process with immigration. *Mathematical notes of the Academy of Sciences of the USSR*, 21(5) :405–411, 1977. 7
- [61] S. Watanabe. On two dimensional markov processes with branching property. *Transactions of the American Mathematical Society*, 136 :447–466, 1969. 4
- [62] S. Watanabe et al. A limit theorem of branching processes and continuous state branching processes. *Journal of Mathematics of Kyoto University*, 8(1) :141–167, 1968. 4
- [63] H.W. Watson and F. Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4 :138–144, 1875. 4
- [64] X. Wu. *Branching processes with biological applications*. PhD thesis, Rice University, 2010. 22
- [65] Y.S. Yang. On branching processes allowing immigration. *Journal of Applied Probability*, 9(1) :24–31, 1972. 6
- [66] A. M. Zubkov. Life-periods of a branching process with immigration. *Theory of Probability & Its Applications*, 17(1) :174–183, 1972. 6