

Modèles aléatoires en Ecologie et Evolution

Sylvie Méléard - Ecole Polytechnique

2009

Table des matières

1	Introduction	7
1.1	Introduction du cours	8
1.2	Importance de la modélisation	9
1.3	De quoi l'on parle...	10
1.3.1	Ecologie et Evolution	10
1.3.2	Génétique des populations	11
1.3.3	Dynamique des populations	12
2	Populations spatiales	13
2.0.4	Marches aléatoires	14
2.0.5	Rappels sur la propriété de Markov	15
2.0.6	Temps de passage en 0	16
2.0.7	Barrières absorbantes	19
2.0.8	Barrières réfléchissantes	21
2.1	Mouvement brownien et diffusions	22
2.1.1	Convergence vers le mouvement brownien	22
2.1.2	Le mouvement brownien	26
2.1.3	La propriété de Markov du mouvement brownien	31
2.2	Martingales et temps d'arrêt	33
2.2.1	Martingales	33
2.2.2	Temps d'arrêt	35
2.2.3	Applications au mouvement brownien	36
2.2.4	Applications aux temps d'atteinte de barrières	39
2.3	Intégrales stochastiques et EDS	40
2.3.1	Intégrales stochastiques	41
2.3.2	Equations différentielles stochastiques (EDS)	43
2.3.3	Propriété de Markov	44

2.3.4	Formule d'Itô	45
2.3.5	Générateur - Lien avec les équations aux dérivées partielles	46
2.3.6	Applications aux temps d'atteinte de barrières	47
3	Dynamique des populations	49
3.1	Processus de population en temps discret	50
3.1.1	Chaînes de Markov de vie et de mort	50
3.1.2	La chaîne de Bienaymé-Galton-Watson	52
3.1.3	Chaîne BGW avec immigration	62
3.1.4	Les probabilités quasi-stationnaires	65
3.1.5	les chaînes densité-dépendantes	71
3.2	Processus markoviens de saut en temps continu	73
3.2.1	Une approche intuitive	73
3.2.2	Processus de Poisson	74
3.2.3	Processus markoviens de saut	82
3.3	Processus de branchement et de naissance et mort en temps continu	91
3.3.1	Processus de branchement en temps continu	91
3.3.2	Cas binaire	95
3.3.3	Extensions	96
3.3.4	Processus de naissance et mort	97
3.3.5	Extinction	98
3.4	Approximations continues : modèles déterministes et stochastiques	101
3.4.1	Approximations déterministes - Equations malthusienne et logistique	101
3.4.2	Approximation stochastique - Stochasticité démographique, Equation de Feller	106
3.5	Populations multi-type	107
3.5.1	Le processus de branchement multi-type en temps discret	107
3.5.2	Les modèles proie-prédateur, systèmes de Lotka-Volterra	113
4	Génétique des populations	115
4.1	Quelques termes de vocabulaire	115
4.2	Le cycle de la reproduction	116
4.3	Introduction	116
4.3.1	Un modèle idéal de population infinie : le modèle de Hardy-Weinberg	116
4.4	Population finie : le modèle de Wright-Fisher	117
4.4.1	Modèle de Wright-Fisher	117
4.4.2	Modèle de Wright-Fisher avec mutation	121

4.4.3	Modèle de Wright-Fisher avec sélection	123
4.5	Modèles démographiques de diffusion	123
4.5.1	Diffusion de Fisher-Wright	123
4.5.2	Diffusion de Fisher-Wright avec mutation et sélection	124
4.5.3	Autre changement d'échelle de temps	126
4.6	La coalescence : description des généalogies	127
4.6.1	Asymptotique quand N tend vers l'infini : le coalescent de Kingman	128
4.6.2	Mutation sur le coalescent	134
4.6.3	Le coalescent avec mutation	135
4.6.4	Urne de Hoppe, restaurant chinois et modèle de Hubbell	137
4.6.5	Le modèle écologique de Hubbell	137
4.6.6	Loi du nombre d'allèles distincts, formule d'Ewens	138
4.6.7	Le point de vue processus de branchement avec immigration	142

Chapitre 1

Introduction

1.1 Introduction du cours

People who have mathematical computation and statistical skills, and establish a collaboration and work on real biological problems, have the chance of doing some very, very significant things for human welfare. (Jaroslav Stark, Imperial College -Science, 2004)

La biologie va d'études très microscopiques, comme la recherche de séquences sur un brin d'ADN, l'étude des échanges moléculaires dans une cellule, l'évolution de tumeurs cancéreuses, l'invasion de parasites dans une cellule, à des problèmes beaucoup plus macroscopiques concernant des comportements de grands groupes d'individus et leurs interactions (extinction de populations, équilibre des éco-systèmes, invasion d'une population par une autre, équilibre proies-prédateurs), ou des problèmes de génétique de populations (recherche d'ancêtres communs à plusieurs individus dans une espèce, phylogénie). A tous les niveaux, l'aléatoire intervient, et certains modèles peuvent aussi bien servir à décrire des problèmes de biologie posés au niveau d'individus microscopiques que de populations macroscopiques.

Le but de ce cours est de montrer l'importance des modèles aléatoires dans la compréhension de la biologie des populations : déplacement de cellules, croissance des bactéries, développement d'une population, réplication de l'ADN, évolution des espèces. L'idée de base est la suivante : même si la population semble présenter un certain nombre de caractéristiques déterministes, elle est composée d'individus dont le comportement est fondamentalement aléatoire et soumis à une grande variabilité. Ainsi, chaque individu se déplace dans une direction différente, chaque bactérie a son propre mécanisme de division cellulaire, chaque réplication de l'ADN peut engendrer une mutation. Pour pouvoir décrire et comprendre comment la population évolue au cours du temps, il faut prendre en compte le comportement de chaque individu. La démarche du probabiliste consiste à déduire du comportement de l'individu, ou de la cellule, ou de la bactérie, des résultats concernant toute la population. Cela permet ainsi, à partir d'une description microscopique très précise d'en déduire des comportements macroscopiques de manière rigoureuse, qu'ils soient déterministes ou aléatoires.

Une caractéristique fondamentale des populations est qu'en général, pour connaître l'évolution de la population dans le futur, il suffit de connaître son état à l'instant présent, et pas nécessairement tout son passé. Cette propriété, appelée *propriété de Markov*, est fondamentale dans la modélisation.

Toute notre démarche va consister à étudier l'évolution de certaines populations au cours du temps. Le temps pourra décrire la succession des différentes générations. Il sera alors discret ($n \in \mathbb{N}$). Nous utiliserons alors la théorie des chaînes de Markov à temps discret, développée dans le cours MAP 432 "Promenade aléatoire". Trois modèles de base classiques illustrent ce propos : les marches aléatoires pour décrire les déplacements spatiaux d'un individu, les processus de Bienaymé-Galton-Watson qui décrivent la dynamique d'une population, le modèle de Wright-Fisher qui décrit une généalogie. Mais le temps peut être aussi le temps physique, à savoir le temps $t \in \mathbb{R}_+$ qui évolue continûment.

Nous serons alors amenés à considérer des processus, à savoir des familles de variables aléatoires indexées par le temps continu, comme les chaînes de Markov à temps continu ou les solutions d'équations différentielles stochastiques. Nous développerons par exemple un certain nombre de résultats sur les processus de branchement en temps continu.

Quand la taille de la population est très grande, il devient difficile de décrire le comportement microscopique de la population (ajout ou retrait de nombreux individus). Nous changerons alors d'échelle de taille de population, et d'échelle de temps, pour nous ramener à des approximations plus facilement manipulables du point de vue mathématique (résultats théoriques et calculs). Dans certaines échelles, nous obtiendrons des approximations déterministes, qui ont été historiquement les premières introduites pour décrire ces évolutions. Dans d'autres échelles nous obtiendrons des approximations aléatoires. C'est ainsi que nous introduirons le mouvement Brownien, un des processus les plus célèbres de l'histoire des probabilités. Des calculs sur ce processus nous permettront d'en déduire un certain nombre d'information sur la population.

Le cours sera développé autour de 3 types de modèles fondamentaux qui seront liés à 3 problématiques différentes concernant les populations : les déplacements spatiaux, la dynamique des populations, la génétique des populations.

1.2 Importance de la modélisation

Insistons sur les points suivants.

Notre but est d'essayer de comprendre, en utilisant un modèle mathématique, l'évolution temporelle, ou 'dynamique', d'un phénomène biologique. Pour pouvoir mettre en évidence ce phénomène, nous serons obligés de simplifier le système biologique pour obtenir un modèle plus facile à étudier mathématiquement. Une difficulté de cette démarche est donc d'obtenir un bon compromis entre le réalisme biologique du modèle et la faisabilité des calculs.

L'approche probabiliste est de ce point de vue plus riche qu'une simple approche déterministe puisqu'elle prend en compte la variabilité individuelle.

L'existence d'un modèle mathématique permettra de pouvoir quantifier numériquement certains phénomènes et de pouvoir prédire certains comportements. (Par exemple montrer qu'une certaine population va s'éteindre et calculer l'espérance du temps d'extinction).

Evidemment, l'on peut toujours se poser la question de la justification du modèle. Une étape ultérieure, dans le cas où l'on peut obtenir des données observées pour le phénomène qui nous intéresse, est de construire des tests statistiques, qui permettront, ou non, de valider le modèle.

Mais terminons par ce qui fait la force du modèle probabiliste. On verra que plusieurs problèmes biologiques très différents (par exemple par les échelles de taille : gènes - cellules

- bactéries - individus - colonies) peuvent avoir des comportements aléatoires de même type et correspondre à des modèles analogues.

Typiquement, introduisons le modèle de naissance et mort le plus simple : Nous définissons

$$X_n = \text{taille de la population à la génération (ou instant) } n,$$

qui évolue de la manière suivante. A chaque instant n , un individu donne naissance ou meurt de manière indépendante de ce qui s'est passé précédemment, avec probabilité $\frac{1}{2}$. Le nombre initial d'individus est X_0 . Alors,

$$X_n = X_0 + \sum_{i=1}^n Z_i,$$

où les variables aléatoires $(Z_i)_i$ sont indépendantes et de même loi $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$. Bien-sûr, comme X_n décrit la taille d'une population et que si il n'y a plus d'individus, on ne peut plus avoir de reproduction, le processus s'arrête quand X_n atteint 0.

Considérons maintenant le modèle de déplacement spatial le plus simple. Un kangourou se déplace sur une route en sautant par sauts de 1 mètre, aléatoirement et indifféremment en avant ou en arrière, à chaque unité de temps. Sa position initiale est X_0 . Alors la position X_n du kangourou au temps n sera

$$X_n = X_0 + \sum_{i=1}^n Z_i,$$

où les variables aléatoires $(Z_i)_i$ sont indépendantes et de même loi $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$. Bien-sûr, ici, il n'y a aucune raison de supposer que le processus s'arrête quand $X_n = 0$. le kangourou va en fait repasser une infinité de fois par 0.

D'un point de vue probabiliste, les deux modèles qui décrivent ces deux phénomènes bien distincts sont très proches.

1.3 De quoi l'on parle...

Le but du cours est d'étudier des modèles probabilistes qui apparaissent en écologie comportementale, en génétique des populations, en dynamique des populations. Tous ces problèmes contribuent à l'étude des écosystèmes et de l'évolution. Quelques définitions sont nécessaires pour préciser de quoi l'on parle.

1.3.1 Ecologie et Evolution

L'**écologie** est l'étude des interactions entre les êtres vivants, entre eux et avec leur milieu, en particulier l'étude de leurs conditions d'existence. L'écologie peut être définie comme

le rapport triangulaire entre les individus d'une espèce, l'activité organisée de cette espèce en sous-populations, et l'environnement de cette activité. L'environnement est à la fois le produit et la condition de cette activité, et donc de la survie de l'espèce. L'écologie étudie donc en particulier la hiérarchie complexe des écosystèmes, les mécanismes biologiques associés à l'extinction des espèces, la dynamique de la biodiversité, l'adaptation des populations.

À son développement contribuent les disciplines plus spécialisées de la dynamique et de la génétique des populations. La dynamique des populations a fait l'objet de modélisations mathématiques depuis longtemps. Les premières études quantitatives de populations reviennent au Révérend Thomas Robert Malthus (1798). Dans son modèle, la croissance de la population est alors géométrique et celle des ressources est arithmétique.

Cette idée nouvelle que la croissance est liée à la limitation des ressources prend toute sa force dans la théorie développée par Darwin, qui défend (en 1859) l'idée fondamentale de sélection naturelle et d'évolution pour les espèces vivantes. On appelle **évolution**, en biologie, la modification des espèces vivantes au cours du temps. Elle est basée sur l'idée simple que les individus les mieux adaptés à leur environnement sont **sélectionnés**, c'est-à-dire sont ceux qui ont les descendanceles plus importantes. L'adaptation découle d'une part des mutations dans les mécanismes de la reproduction, qui créent de la variabilité dans les caractéristiques des individus, et d'autre part de la compétition entre ceux-ci, due au partage des ressources, qui va permettre de sélectionner les plus aptes à les utiliser.

Pour comprendre l'évolution des espèces, il faut être capable de comprendre les mécanismes internes des généalogies à l'intérieur d'une espèce. C'est un des buts de la génétique des populations, qui constitue l'approche mathématique la plus ancienne de modélisation de l'évolution.

1.3.2 Génétique des populations

La génétique des populations est l'étude de la fluctuation des allèles (les différentes versions d'un gène) au cours du temps dans les populations d'individus d'une même espèce, sous l'influence de

- la dérive génétique (la variabilité due au hasard des événements individuels de naissance et de mort),
- les mutations,
- la sélection naturelle,
- les migrations.

Elle a été initiée par les biologistes Fisher, Haldane, Wright entre 1920 et 1940. Elle se développe à partir des principes fondamentaux de la génétique mendélienne et de la théorie darwinienne de l'évolution. Elle a de grandes applications en épidémiologie (permet d'étudier la transmission des maladies génétiques), en agronomie (des programmes de sélection modifient le patrimoine génétique de certains organismes pour créer des races ou variétés plus performantes... OGM). Elle permet également de comprendre les mécanismes

de conservation et de disparition des populations et des espèces.

1.3.3 Dynamique des populations

A vital next step will be to promote the training of scientists with expertise in both mathematics and biology. (Science 2003).

La dynamique des populations étudie la répartition et le développement quantitatif de populations d'individus, asexués ou sexués. Elle s'intéresse à

- des mécanismes d'auto-régulation des populations : les paramètres de croissance de la population dépendent de la taille de la population,
- des modèles de prédation, et les interactions entre prédateur et proie,
- des modèles de coopération-compétition.

Chapitre 2

Populations spatiales

The mathematics is not there till we put it there. Sir Arthur Eddington (1882 - 1944),
The Philosophy of Physical Science.

Nous allons commencer par modéliser la dynamique spatiale d'une population. Bien sûr, la **dispersion** d'une population est fondamentalement liée à la dynamique de la population et l'idéal serait de prendre en compte tous les facteurs. Mais comme les modèles spatiaux sont dans une première approximation plus simples du point de vue mathématique que les modèles de dynamique des populations, c'est par eux que nous allons commencer ce cours.

La dispersion des individus d'une population peut être liée

- à la recherche de nourriture
- à une recherche d'extension territoriale,
- au besoin de se déplacer dans un endroit mieux adapté qui maximise le taux de croissance et minimise la probabilité d'extinction de cette population.

Les problèmes de colonisation de nouveaux territoires, d'invasion et de structure spatiale de la biodiversité sont fondamentaux. On peut citer par exemple les études faites sur

- la progression des lapins en Europe centrale,
- l'invasion du cerf-rouge en Nouvelle-Zélande (voir Renshaw [19]),
- la vitesse d'invasion surprenante des crapauds-buffles en Australie (voir [18]).

Soit à cause de frontières naturelles (La mer pour une île, une chaîne de montagne), ou de limitation de la zone de ressources, les populations se déplacent souvent dans des domaines fermés et le comportement au bord du domaine est important. La barrière est dite absorbante si les animaux meurent ou restent bloqués à son contact. (Un poisson capturé au bord d'un bassin). Elle est dite réfléchissante si l'individu rebrousse chemin à son contact (le poisson au bord de l'aquarium).

Nous allons dans ce premier chapitre étudier les dynamiques d'une population évoluant

soit dans un espace qui pourra être \mathbb{R} , \mathbb{R}^2 ou \mathbb{R}^3 , soit dans un sous-domaine de l'espace. Pour l'instant nous ne tiendrons pas compte de la démographie de cette population (naissances et morts), mais nous ferons toutefois allusion à plusieurs occasions à l'analogie entre dynamiques spatiale et démographique développée dans l'introduction.

2.0.4 Marches aléatoires

Dans ce chapitre nous utiliserons principalement les outils de chaînes de Markov à temps discret. Nous renvoyons au cours MAP 432 "Promenade aléatoire" ([2]).

Le modèle aléatoire le plus simple pour décrire le déplacement au hasard d'un individu est celui d'une marche aléatoire simple. L'espace est modélisé par le réseau \mathbb{Z}^d et l'individu, à partir d'un point $x = (x_1, \dots, x_d)$ peut aller vers chacun de ses voisins immédiats avec la même probabilité $\frac{1}{2d}$, si la marche est symétrique, ou avec des probabilités différentes (mais de somme 1) si la marche est non symétrique. Les déplacements successifs de l'individu sont indépendants les uns des autres. Il est alors clair que le processus qui décrit la position de l'individu au cours du temps est markovien.

Ainsi, si l'espace est \mathbb{Z} , modélisant par exemple une rivière ou une route, ou bien la projection sur une dimension d'un déplacement dans l'espace, la marche aléatoire simple sera une chaîne de Markov de matrice de transition (P_{ij}) vérifiant

$$P_{i,i+1} = p ; P_{i,i-1} = 1 - p = q.$$

Si $p = \frac{1}{2}$, la marche est symétrique.

Si X_n est la position à l'instant n d'un individu dont le déplacement suit une marche aléatoire simple, on a

$$X_n = X_0 + \sum_{i=1}^n Z_i,$$

où les variables aléatoires Z_i sont indépendantes et équidistribuées de loi chargeant $\{-1, 1\}$ avec probabilités $1 - p$ et p . Si $0 < p < 1$, alors tous les états communiquent. Ainsi, si la marche passe par un point précis une infinité de fois (ce point est appelé point récurrent), elle passera par tous les points une infinité de fois.

On peut généraliser ce modèle en supposant que les variables aléatoires Z_n sont indépendantes et de même loi, et l'on note

$$E(Z_n) = m ; Var(Z_n) = \sigma^2. \quad (2.1)$$

Remarquons que ces marches aléatoires sont des exemples simples de chaînes de Markov, et vérifient la propriété de Markov forte (voir cours MAP 432).

Nous allons nous intéresser tout d'abord aux marches aléatoires en dimension 1 qui, comme nous l'avons vu dans l'introduction, décrivent non seulement des déplacements spatiaux, mais aussi des dynamiques de tailles de populations.

2.0.5 Rappels sur la propriété de Markov

La propriété de Markov décrit une propriété satisfaite par de nombreux phénomènes aléatoires, pour lesquels l'évolution future ne dépend du passé qu'à travers sa valeur au temps présent. Pour de plus amples détails, nous renvoyons au cours MAP 432.

Les marches aléatoires introduites ci-dessus vérifient cette propriété et sont un cas particulier simple de chaînes de Markov. Dans la suite du cours, nous verrons un grand nombre de processus vérifiant la propriété de Markov.

Proposition 2.0.1 *La marche aléatoire $(X_n)_{n \in \mathbb{N}}$ à valeurs dans \mathbb{Z} satisfait la **propriété de Markov** si pour tous $n \in \mathbb{N}$, $j, i_0, \dots, i_n \in \mathbb{Z}$, tels que $\mathbb{P}(X_n = i_n, \dots, X_0 = i_0) > 0$, on a*

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i_n). \quad (2.2)$$

La loi conditionnelle de X_{n+1} sachant X_0, \dots, X_n est égale à sa loi conditionnelle sachant X_n .

Nous allons introduire la suite de tribus $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ engendrées par la chaîne. Ces tribus sont croissantes (pour l'inclusion). La suite $(\mathcal{F}_n)_n$ s'appelle la filtration engendrée par la marche aléatoire. Remarquons que \mathcal{F}_n décrit toute l'information donnée par le processus jusqu'à l'instant n . Il est naturel que cette information croisse avec n . Il est important de connaître la marche aléatoire aux temps discrets n , mais certains temps aléatoires vont aussi s'avérer fondamentaux, en particulier ceux dont la connaissance est liée à celle du processus. Pour la marche aléatoire $(X_n)_n$, les temps successifs de passage de la marche en l'état i vont être essentiels. On les définit ainsi :

$$T_i = \inf(n \geq 1 : X_n = i) \quad (2.3)$$

$$T_i^1 = T_i, \quad T_i^{k+1} = \inf(m > T_i^k : X_m = i) \quad (2.4)$$

avec la convention habituelle que l'infimum de l'ensemble vide vaut $+\infty$. La suite $(T_i^k)_k$ décrit les temps de passage successifs de la chaîne en l'état i .

Remarquons que l'observation de X_0, \dots, X_n , c'est-à-dire de la chaîne jusqu'à l'instant n , permet de décider si T_i vaut n ou non. Ainsi, par la définition de \mathcal{F}_n , on aura

$$\{T_i = n\} \in \mathcal{F}_n.$$

Plus généralement, on va s'intéresser à tous les temps aléatoires qui vérifient cette propriété.

Définition 2.0.2 *On appelle **temps d'arrêt** une variable aléatoire T à valeurs dans $\mathbb{N} \cup \{+\infty\}$ qui vérifie que pour tout $n \in \mathbb{N}$,*

$$\{T = n\} \in \mathcal{F}_n. \quad (2.5)$$

Remarque : T_i est bien un temps d'arrêt, puisque

$$\{T_i = n\} = \{X_1 \neq i\} \cap \dots \cap \{X_{n-1} \neq i\} \cap \{X_n = i\}.$$

Exemple 2.0.3 Pour tout $A \subset E$, le temps d'atteinte de A défini par

$$T_A = \inf(n \geq 1 : X_n \in A)$$

est un temps d'arrêt.

Par contre, l'instant de dernier passage en A

$$L_A = \sup(n \geq 1 : X_n \in A)$$

n'est pas un temps d'arrêt. Pour connaître L_A , il faut avoir de l'information sur les états ultérieurs de la chaîne.

On peut montrer que la marche aléatoire satisfait encore la propriété de Markov si l'on considère sa position à un temps d'arrêt. Cette propriété s'appelle la propriété de Markov forte.

Théorème 2.0.4 (Propriété de Markov forte). Soit $(X_n)_n$ une marche aléatoire et T un temps d'arrêt. Conditionnellement à X_T sur $\{T < \infty\}$, la suite $(X_{T+n})_{n \geq 1}$ est une chaîne de Markov indépendante de la trajectoire de $(X_n)_n$ jusqu'au temps T .

2.0.6 Temps de passage en 0

Nous allons étudier la marche aléatoire simple partant de 0. On supposera donc ici que $X_0 = 0$. Nous allons nous intéresser à la suite des instants aléatoires où la particule se retrouve à l'origine. Pour décrire ces instants, il suffit de considérer l'instant de premier retour à l'origine, puisque les instants suivants sont des sommes de copies indépendantes de celui-ci par la propriété de Markov forte. Soit donc

$$T_0 = \min\{n \geq 1 : X_n = 0\} \in \mathbb{N}^* \cup \{\infty\}$$

l'instant de premier retour en 0. (On a posé $\min\{\emptyset\} = 0$). T_0 est un temps d'arrêt pour la filtration $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ engendrée par la chaîne. Nous souhaitons connaître la loi de T_0 . Pour cela, introduisons

$$f(n) = \mathbb{P}(T_0 = n) = \mathbb{P}(X_1 \neq 0, \dots, X_{n-1} \neq 0, X_n = 0),$$

pour $n \geq 1$. Attention, il se peut que $\mathbb{P}(T_0 = +\infty) > 0$ (si le processus ne revient jamais en 0), auquel cas la fonction génératrice F_0 de T_0 , définie pour $s \in [0, 1]$ par

$$F(s) = \mathbb{E}(s^{T_0}) = \sum_{n=1}^{\infty} f(n) s^n,$$

vérifie $F(1) = 1 - \mathbb{P}(T_0 = \infty) = \mathbb{P}(T_0 < \infty)$. Nous allons calculer F , à l'aide de la fonction auxiliaire Q définie pour $s \in [0, 1]$ par

$$Q(s) = \sum_{n=0}^{\infty} \mathbb{P}(X_n = 0) s^n,$$

que l'on calcule plus aisément.

Proposition 2.0.5 *On a les trois égalités suivantes : Pour tout $s \in [0, 1]$,*

(i) $Q(s) = 1 + Q(s) F(s)$

(ii) $Q(s) = (1 - 4pq s^2)^{-1/2}$

(iii) $F(s) = 1 - (1 - 4pq s^2)^{1/2}$.

Preuve. D'après la formule de Bayes,

$$\mathbb{P}(X_n = 0) = \sum_{k=1}^n \mathbb{P}(X_n = 0 | T_0 = k) \mathbb{P}(T_0 = k).$$

Il est facile de voir que, puisque les variables aléatoires Z_i , ($i \geq 1$) sont indépendantes et de même loi, l'on a $\mathbb{P}(X_n = 0 | T_0 = k) = \mathbb{P}(X_n - X_k = 0 | T_0 = k) = \mathbb{P}(X_{n-k} = 0)$, d'où $a_n = \mathbb{P}(X_n = 0)$ est solution de

$$a_n = \sum_{k=1}^n a_{n-k} f(k),$$

et $a_0 = 1$. En multipliant par s^n et en sommant en n , puis en utilisant le théorème de Fubini, nous obtenons

$$\begin{aligned} Q(s) &= \sum_{n \geq 0} a_n s^n = 1 + \sum_{n \geq 1} \sum_{k=1}^n a_{n-k} s^{n-k} f(k) s^k \\ &= 1 + \sum_{k=1}^{\infty} f(k) s^k \sum_{n=k}^{+\infty} a_{n-k} s^{n-k} = 1 + F(s) Q(s), \end{aligned}$$

ce qui démontre (i).

Notons que $\mathbb{P}(X_n = 0) = \binom{n}{n/2} (pq)^{n/2}$ lorsque n est pair, et $\mathbb{P}(X_n = 0) = 0$ si n est impair, puisque $(X_n + n)/2$ suit la loi binomiale $B(n, p)$. On a alors que

$$\begin{aligned} Q(s) &= \sum_{n \geq 0} \binom{2n}{n} (pq)^n s^{2n} = \sum_{n \geq 0} \frac{2^n (2n-1)(2n-3) \cdots 1}{n!} (pq s^2)^n \\ &= \sum_{n \geq 0} \frac{(2n-1)(2n-3) \cdots 1}{2^n n!} (4pq s^2)^n. \end{aligned}$$

Nous pouvons identifier le dernier terme en le comparant au développement en série entière

$$(1+u)^{-\frac{1}{2}} = 1 - \frac{1}{2}u + \frac{(-\frac{1}{2})(-\frac{1}{2}-1)}{2!}u^2 + \dots + \frac{(-\frac{1}{2})(-\frac{1}{2}-1)\dots(-\frac{1}{2}-n+1)}{n!}u^n + \dots$$

(pour $|u| < 1$). Nous en déduisons alors (ii).

Enfin, (iii) découle immédiatement des deux premiers points. \square

Corollaire 2.0.6 *La probabilité que la particule retourne au moins une fois à l'origine est*

$$\mathbb{P}(T_0 < \infty) = F(1) = 1 - |2p - 1| = \begin{cases} 2(1-p) & \text{si } p > 1/2 \\ 2p & \text{si } p < 1/2 \\ 1 & \text{si } p = 1/2 \end{cases}$$

Si $p = \frac{1}{2}$, le temps moyen de premier retour est tel que

$$\mathbb{E}(T_0) = \lim_{s \nearrow 1} F'(s) = +\infty.$$

La marche aléatoire est dite **récurrente** si la particule finit par revenir à son point de départ avec probabilité 1, elle est dite **transiente** sinon.

D'après le corollaire, il y a récurrence si et seulement si $p = \frac{1}{2}$. Ceci est conforme à la loi des grands nombres, qui entraîne ici que $\mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = 2p - 1\right) = 1$.

Notons que lorsque $p = 1/2$, $T_0 < \infty$ avec probabilité un, mais le retour à 0 s'effectue lentement puisque le temps moyen de retour à 0 est infini.

Preuve. On a

$$\begin{aligned} \mathbb{P}(T_0 < \infty) &= \lim_{z \nearrow 1} F(s) = 1 - (1 - 4pq)^{1/2} \\ &= 1 - |2p - 1| \end{aligned}$$

d'après (iii). Enfin, lorsque $p = 1/2$, $F(s) = 1 - (1 - s^2)^{1/2}$ et $\mathbb{E}_0(T_0) = \lim_{s \nearrow 1} F'(s) = +\infty$. \square

Remarquons que dans le cas général (2.1), nous connaissons par le théorème de la limite centrale le comportement asymptotique du processus X_n . En effet, on sait que $\frac{X_n - nm}{\sigma\sqrt{n}}$ converge en loi, quand n tend vers l'infini, vers une loi normale centrée réduite. Ainsi,

$$X_n \approx nm + \sqrt{n}N(0, \sigma^2) + o(\sqrt{n})$$

quand $n \rightarrow \infty$. Nous savons également par la loi des grands nombres que $\frac{X_n}{n}$ converge presque-sûrement vers m quand n tend vers l'infini. Nous en déduisons que si $m \neq 0$, la

chaîne est forcément transiente : à partir d'un certain rang, elle ne pourra plus repasser par 0. Le cas où $m = 0$ est beaucoup plus délicat. On peut s'en rendre compte en étudiant les marches aléatoires symétriques. Nous avons étudié le cas de la dimension 1, mais on montre en fait que les propriétés de transience et de récurrence de la marche aléatoire symétrique sont liées à la dimension de l'espace. Nous avons le théorème (difficile) suivant.

Théorème 2.0.7 (voir MAP 432) *La marche aléatoire simple symétrique est récurrente si la dimension d est égale à 1 ou 2. Si $d \geq 3$, la marche aléatoire est transiente.*

Rappelons également le résultat suivant, qui pour une marche aléatoire transiente donne le nombre moyen de retours à 0. La démonstration utilise la propriété de Markov

Proposition 2.0.8 (voir MAP 432) *On appelle N_0 le nombre de retours à l'état 0. Si (X_n) est une marche aléatoire récurrente, alors*

$$\mathbb{P}(N_0 = +\infty) = 1.$$

Si (X_n) est une marche aléatoire transiente et si $\pi_0 = \mathbb{P}(T_0 < \infty)$. Alors N_0 suit une loi géométrique de paramètre π_0 :

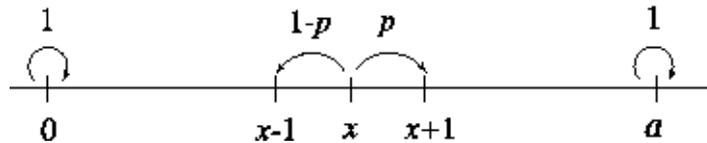
$$\mathbb{P}(N_0 = k) = \pi_0^k (1 - \pi_0).$$

Dans ce cas, le nombre moyen de retours à 0 vaut

$$\mathbb{E}(N_0) = \frac{\pi_0}{1 - \pi_0} < \infty.$$

2.0.7 Barrières absorbantes

Considérons une marche aléatoire en dimension 1 avec deux **barrières absorbantes** aux points 0 et $a \in \mathbb{N}^*$. Pour une population spatiale, telle une population de poissons dans une rivière, on peut imaginer deux filets de pêche en positions 0 et a . Si le poisson atteint un des filets, il est capturé et reste immobile. On dit que la marche est absorbée en 0 et en a . Il est commode de représenter la dynamique de cette marche par le diagramme suivant qui indique la probabilité des différents sauts possibles.



On peut également s'intéresser à ce type de modèles pour des dynamiques de populations. Imaginons que X_n modélise la dynamique d'une taille de population. Celle-ci croît de 1 avec probabilité p et décroît de 1 avec probabilité $1-p$. Le passage en 0 signifie l'extinction de la population. Une fois éteinte, la taille de la population reste égale à 0. On a donc une marche aléatoire à valeurs dans \mathbb{N} , que l'on dit **absorbée en 0**. Nous allons voir que suivant le rapport entre les probabilités de naissance et mort (p et $1-p$), la population va s'éteindre ou non.

Probabilité d'absorption en 0 ou d'extinction

Notons \mathbb{P}_k la probabilité conditionnée au fait que $X_0 = k$. Notons G l'événement "la marche atteint 0". Nous allons déterminer la probabilité $\mu_k = \mathbb{P}_k(G)$ de G sachant que la position initiale de l'individu est k . On a immédiatement

$$\mathbb{P}_k(G) = \mathbb{P}_k(G | Z_1 = +1) \mathbb{P}(Z_1 = +1) + \mathbb{P}_k(G | Z_1 = -1) \mathbb{P}(Z_1 = -1), \quad (2.6)$$

d'où

$$p \mu_{k+1} + (1-p) \mu_{k-1} - \mu_k = 0 \quad (2.7)$$

pour $k \in \{1, \dots, a-1\}$, avec les conditions aux limites $\mu_0 = 1$, $\mu_a = 0$. L'équation (2.7) est une récurrence linéaire, pour laquelle on cherche d'abord les solutions de la forme $\mu_k = r^k$: r doit satisfaire l'équation caractéristique

$$p r^2 - r + (1-p) = 0,$$

dont les solutions sont $r_1 = 1$, $r_2 = (1-p)/p$. Deux cas se présentent alors.

- $p \neq 1/2$ (**marche asymétrique**). Les deux racines sont différentes, les solutions de (2.7) sont de la forme $\mu_k = \alpha r_1^k + \beta r_2^k$, et on détermine α et β par les conditions aux limites $\mu_0 = 1$, $\mu_a = 0$. On obtient

$$\mathbb{P}_k(G) = \frac{\left(\frac{1-p}{p}\right)^a - \left(\frac{1-p}{p}\right)^k}{\left(\frac{1-p}{p}\right)^a - 1}, \quad 0 \leq k < a \quad (2.8)$$

Procédant de même pour l'événement D : "la marche aléatoire atteint a ", on constate que $\mathbb{P}_k(D)$ satisfait (2.7), avec d'autres conditions aux limites ($\mu_0 = 0$, $\mu_a = 1$). On trouve alors que $\mathbb{P}_k(D) = 1 - \mathbb{P}_k(G)$, si bien que l'individu est piégé en un temps fini, avec probabilité 1.

- $p = 1/2$ (**marche symétrique**). Alors $r_1 = r_2 = 1$, les solutions de (2.7) sont de la forme $\mu_k = \alpha + \beta k$, et tenant compte des conditions limites on obtient

$$\mathbb{P}_k(G) = 1 - \frac{k}{a}, \quad \mathbb{P}_k(D) = \frac{k}{a}, \quad (2.9)$$

Ici encore le poisson finit par être piégé.

Durée moyenne avant d'être piégé

Déterminons $V_k = \mathbb{E}_k(T)$ où T est le temps d'atteinte d'une des barrières, défini par

$$T = \min\{n \geq 0 : X_n = 0 \text{ ou } a\} = T_0 \wedge T_a.$$

Ici \mathbb{E}_k désigne l'espérance sous \mathbb{P}_k . Par la propriété de Markov, en conditionnant au comportement de la chaîne au temps 1, on obtient que pour $1 \leq k \leq a-1$,

$$\mathbb{E}_k(T) = 1 + \mathbb{P}(Z_1 = +1) \mathbb{E}_{k+1}(T) + \mathbb{P}(Z_1 = -1) \mathbb{E}_{k-1}(T).$$

On obtient ainsi l'équation de récurrence linéaire avec second membre

$$pV_{k+1} + (1-p)V_{k-1} - V_k = -1 \quad (2.10)$$

pour $1 \leq k \leq a-1$, avec $V_0 = V_a = 0$. On résout cette équation en trouvant les solutions générales et une solution particulière. (L'équation caractéristique est la même que dans les calculs précédents).

$$\begin{aligned} \mathbb{E}_k(\nu) &= \frac{1}{1-2p} \left(k - a \frac{1 - \left(\frac{1-p}{p}\right)^k}{1 - \left(\frac{1-p}{p}\right)^a} \right) && \text{si } p \neq 1/2, \\ &= k(a-k) && \text{si } p = 1/2. \end{aligned}$$

Extinction de population

Remarquons que si $a \rightarrow \infty$, et si $1-p < p$, alors la probabilité d'arriver à 0 en partant de k vaut $\left(\frac{1-p}{p}\right)^k$. En revanche, si $1-p \geq p$, cette probabilité vaut 1 pour tout k . L'interprétation de ce résultat en terme de dynamique de population est claire : si la probabilité de naissance p est supérieure à la probabilité de mort $1-p$, il y a une probabilité strictement positive de ne pas d'éteindre, mais dans le cas contraire, la probabilité d'extinction est 1.

2.0.8 Barrières réfléchissantes

Nous nous plaçons toujours dans le cas d'une marche aléatoire simple en dimension 1. Par exemple, si elle modélise le comportement d'un poisson dans une rivière, supposons maintenant que lorsque le poisson touche une des frontières 0 ou a de son domaine, il repart instantanément dans l'autre sens avec une certaine probabilité p (resp. $1-p$) et reste capturé sinon. On peut alors procéder comme précédemment, sachant que quand le poisson est en 0 (resp. en a), il a une probabilité $1-p$ (resp. p) d'y rester, et p (resp. $1-p$) de repartir en 1 (resp. en $a-1$).

On peut imaginer que le poisson se déplace indéfiniment entre les barrières avec probabilité positive. Il est alors naturel de chercher le comportement de la marche aléatoire quand $n \rightarrow \infty$. Puisque (X_n) ne prend qu'un nombre fini de valeurs, nous savons qu'il existe au moins une probabilité invariante $\pi = (\pi_i)_i$ (cf. MAP 432). Nécessairement, en prenant en compte le déplacement antérieur des individus, nous obtenons que π satisfait l'équation suivante

$$\pi_i = p\pi_{i-1} + (1-p)\pi_{i+1}, \quad 1 \leq i \leq a-1$$

avec les conditions aux bords

$$\pi_0 = (1-p)\pi_0 + (1-p)\pi_1 ; \pi_a = p\pi_{a-1} + p\pi_a.$$

On peut facilement résoudre l'équation et montrer que nécessairement

$$\pi_i = \left(\frac{p}{1-p}\right)^i \frac{1 - \left(\frac{p}{1-p}\right)^{a+1}}{1 - \left(\frac{p}{1-p}\right)^{a+1}}, \quad \forall i = 0, \dots, a.$$

On a donc unicité de la probabilité invariante.

Si $p = \frac{1}{2}$, on a $\pi_i = \frac{1}{a+1}$ et la distribution uniforme est invariante sur $\{0, \dots, a\}$, ce qui est un résultat très intuitif. Si en revanche, $1-p > p$ (resp. $1-p < p$), alors π_i décroît géométriquement vite depuis la barrière 0 (resp. depuis a).

Les exemples que nous venons de développer en dimension 1 sont très simples, mais montrent toutefois que les calculs donnant les probabilités d'atteinte de barrières, de durée moyenne avant d'être piégé ou de probabilité invariante, deviennent vite compliqués. On peut imaginer que des dynamiques plus complexes (dimension plus grande, forme compliquée du domaine spatial, ...) peuvent amener des difficultés techniques énormes. Nous allons maintenant introduire une approximation de ces marches aléatoires dans des échelles spatiale et temporelle différentes, qui permettra plus facilement de faire des calculs. Cette approximation est un processus à trajectoires continues fondamental dans la modélisation probabiliste : **le mouvement brownien**.

2.1 Mouvement brownien et diffusions

Pour tout ce chapitre (mouvement brownien et calcul stochastique), nous renvoyons prioritairement au polycopié de El Karoui-Gobet [7]

2.1.1 Convergence vers le mouvement brownien

En reprenant les notations du chapitre précédent, nous pouvons remarquer qu'une marche aléatoire vérifie les deux propriétés fondamentales suivantes :

- Pour tous n et k , $X_{n+k} - X_n$ est indépendant de X_n
- Pour tous n et k , $X_{n+k} - X_n$ a même loi que X_k .

On dit que les accroissements sont indépendants et stationnaires.

Supposons que la marche aléatoire est **simple symétrique** sur \mathbb{Z} et que $X_0 = 0$. Elle saute d'une amplitude de l'ordre d'une unité d'espace pendant chaque unité de temps. Mais ce type de modèle est inadéquat dès lors que l'on considère des mouvements très

erratiques et qui semblent varier presque continûment : évolution de molécules, de planctons, d'insectes, déplacement de particules de pollen à la surface de l'eau. Dans de telles situations, les échelles de taille sont très petites et les changements de position très rapides.

Nous allons donner un changement d'échelles d'espace et de temps qui va permettre, à partir d'une marche aléatoire, de définir un nouvel objet mathématique décrivant ce type de comportement. La bonne échelle va nous être donnée par le théorème de la limite centrale.

Fixons un intervalle de temps $[0, t]$. On va supposer qu'il y a un saut de la marche aléatoire tous les temps $\frac{1}{n}$. Il y aura donc $[nt]$ sauts sur l'intervalle $[0, t]$, (où $[nt]$ désigne la partie entière de nt). On étudie alors la marche aléatoire

$$X_{[nt]} = \sum_{k=1}^{[nt]} Z_k$$

et notre but est d'en connaître le comportement asymptotique, quand n tend vers l'infini. Remarquons que comme l'individu saute tous les temps $\frac{1}{n}$, cette limite rend les changements de position très rapides.

Comme l'espérance de $X_{[nt]}$ est nulle, et que sa variance vaut $[nt]$, la renormalisation naturelle, pour avoir un objet probabiliste convergeant quand n tend vers l'infini, est une échelle spatiale d'ordre $\frac{1}{\sqrt{n}}$. Cette approche intuitive est confirmée par le théorème de la limite centrale. En effet, les variables aléatoires Z_k sont indépendantes et de même loi centrée réduite. Donc quand t est fixé,

$$\frac{X_{[nt]}}{\sqrt{[nt]}} \xrightarrow[n \rightarrow \infty]{\text{en loi}} N(0, 1), \quad (2.11)$$

où $N(0, 1)$ désigne une variable aléatoire de loi normale centrée réduite. Nous pouvons alors écrire

$$\frac{X_{[nt]}}{\sqrt{[nt]}} = \frac{X_{[nt]}}{\sqrt{n}} \frac{\sqrt{n}}{\sqrt{[nt]}}$$

et donc en utilisant que $\frac{\sqrt{n}}{\sqrt{[nt]}}$ converge vers $\frac{1}{\sqrt{t}}$ quand n tend vers l'infini, nous en déduisons la

Proposition 2.1.1

$$X_t^n = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} Z_k = \frac{X_{[nt]}}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\text{en loi}} N(0, t), \quad (2.12)$$

où $N(0, t)$ désigne une variable aléatoire de loi normale centrée et de variance t .

Remarque 2.1.2 1) Si nous avons supposé que les variables aléatoires Z_k étaient de moyenne $m \neq 0$, alors nous aurions un comportement asymptotique différent. En effet, en utilisant la loi des grands nombres,

$$\frac{X_{[nt]}}{n} \xrightarrow[n \rightarrow \infty]{p.s.} mt.$$

Ainsi, les sauts sont d'amplitude $\frac{1}{n}$ et la limite est déterministe.

2) Si nous considérons le processus W_t^n défini comme étant la marche aléatoire simple σX_t^n sur $\frac{\sigma}{\sqrt{n}}\mathbb{Z}$, chaque pas étant perturbé par $\frac{\mu}{n}$, alors nous aurons

$$W_t^n = \sum_{k=1}^{[nt]} \left(\frac{\sigma}{\sqrt{n}} Z_k + \frac{\mu}{n} \right) \xrightarrow[n \rightarrow \infty]{en\ loi} \sigma N(0, t) + \mu t. \quad (2.13)$$

La limite est alors la somme d'une variable aléatoire normale et d'un terme déterministe.

Nous allons maintenant étudier la dynamique du processus $t \rightarrow X_t^n$. Ce processus saute d'une amplitude $\frac{1}{\sqrt{n}}$ tous les temps $\frac{1}{n}$. Nous pouvons aussi en considérer une approximation linéaire à trajectoires continues en posant

$$Y_t^n = \frac{X_{[nt]}}{\sqrt{n}} + \frac{nt - [nt]}{\sqrt{n}} Z_{[nt]+1}.$$

Alors, puisque $\frac{nt - [nt]}{\sqrt{n}} \leq \frac{1}{\sqrt{n}}$, on a que $\mathbb{E} \left(\left(\frac{nt - [nt]}{\sqrt{n}} Z_{[nt]+1} \right)^2 \right)$ tend vers 0 quand n tend vers l'infini. Ainsi, la suite de variables aléatoires $\left(\frac{nt - [nt]}{\sqrt{n}} Z_{[nt]+1} \right)_n$ converge en moyenne quadratique et donc en loi vers 0. Par ailleurs, pour chaque n , les variables $\frac{X_{[nt]}}{\sqrt{n}}$ et $\frac{nt - [nt]}{\sqrt{n}} Z_{[nt]+1}$ sont indépendantes. Nous pouvons alors utiliser le théorème de Lévy (voir MAP 311) pour étudier la convergence en loi de la suite Y_t^n . Soit $u \in \mathbb{R}$. On a

$$\mathbb{E} \left(e^{iuY_t^n} \right) = \mathbb{E} \left(e^{iu \frac{X_{[nt]}}{\sqrt{n}}} \right) \mathbb{E} \left(e^{iu \frac{nt - [nt]}{\sqrt{n}} Z_{[nt]+1}} \right) \rightarrow e^{-\frac{tu^2}{2}} \times 1.$$

Cela entraîne donc que pour chaque t fixé, la suite de variables aléatoires $(Y_t^n)_n$ converge en loi vers $N(0, t)$. Là-encore, nous pouvons nous intéresser au comportement du processus continu $t \rightarrow Y_t^n$. Nous allons montrer que les processus $t \rightarrow X_t^n$ et $t \rightarrow Y_t^n$ convergent vers la même limite quand n tend vers l'infini. Cette limite est illustrée par les simulations données dans les pages suivantes en dimension 1 et dimension 2.

On observe que la limite est un processus à trajectoires continues en temps mais qu'il est très fortement irrégulier. Nous allons mettre en évidence mathématiquement ce processus.

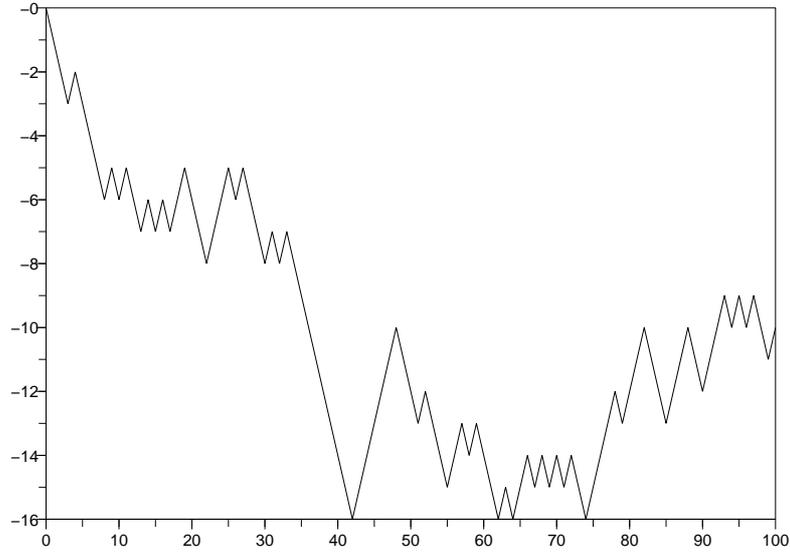


FIG. 2.1 – marche aléatoire en dimension 1

La continuité de la limite de $(t \rightarrow X_t^n)_n$ n'est pas surprenante puisque l'amplitude des sauts de X_t^n vaut $\frac{1}{\sqrt{n}}$ et tend donc vers 0 quand n tend vers l'infini.

Remarquons tout d'abord que nous avons mieux que la convergence "t par t" de $(X_t^n)_n$. En effet, considérons une suite finie de temps $t_1 < t_2 < \dots < t_p$. Nous savons que les variables

$$X_{t_1}^n, X_{t_2}^n - X_{t_1}^n, \dots, X_{t_p}^n - X_{t_{p-1}}^n$$

sont indépendantes, et de plus,

$$\begin{aligned} X_{t_j}^n - X_{t_{j-1}}^n &= \frac{1}{\sqrt{n}} (Z_{[nt_{j-1}]+1} + \dots + Z_{[nt_j]}) \\ &= \frac{\sqrt{[nt_j] - [nt_{j-1}]}}{\sqrt{n}} \times \frac{Z_{[nt_{j-1}]+1} + \dots + Z_{[nt_j]}}{\sqrt{[nt_j] - [nt_{j-1}]}}. \end{aligned}$$

Le premier terme du membre de droite tend vers $\sqrt{t_j - t_{j-1}}$ et le deuxième vers une variable aléatoire normale centrée réduite. Comme précédemment, et puisque les coordonnées sont indépendantes, nous en déduisons que

$$(X_{t_1}^n, X_{t_2}^n - X_{t_1}^n, \dots, X_{t_p}^n - X_{t_{p-1}}^n) \xrightarrow[n \rightarrow \infty]{\text{en loi}} (N(0, t_1), N(0, t_2 - t_1), \dots, N(0, t_p - t_{p-1})),$$

quand n tend vers l'infini. Mais alors, nous pouvons écrire

$$X_{t_2}^n = X_{t_1}^n + (X_{t_2}^n - X_{t_1}^n); \dots; X_{t_p}^n = X_{t_1}^n + \dots + (X_{t_p}^n - X_{t_{p-1}}^n)$$

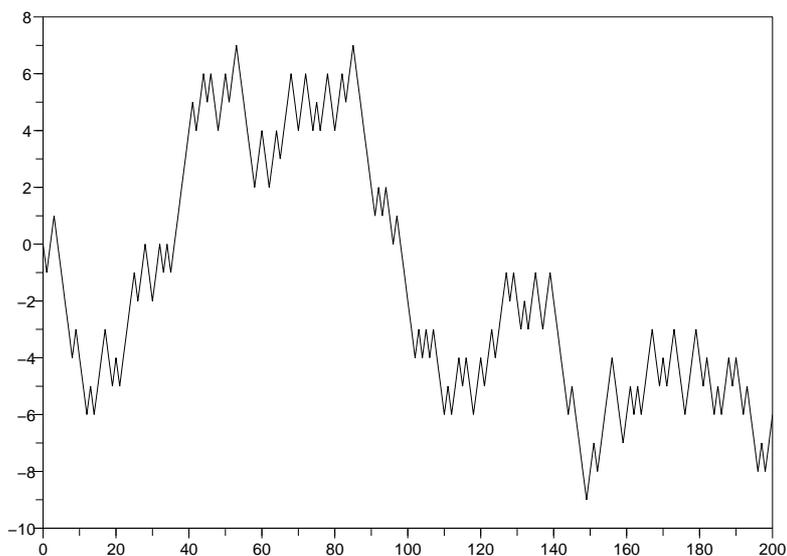


FIG. 2.2 – marche aléatoire en dimension 1

et en déduire que quand n tend vers l'infini,

$$(X_{t_1}^n, X_{t_2}^n, \dots, X_{t_p}^n) \xrightarrow[n \rightarrow \infty]{\text{en loi}} (B_{t_1}, B_{t_2}, \dots, B_{t_p}),$$

où $B_t - B_s$ est pour $s < t$ une variable aléatoire de loi normale centrée de variance $t - s$ indépendante de B_s .

Nous avons donc montré que chaque marginale fini-dimensionnelle $(X_{t_1}^n, X_{t_2}^n, \dots, X_{t_p}^n)$ du processus (X_t^n) converge en loi vers la marginale correspondante $(B_{t_1}, B_{t_2}, \dots, B_{t_p})$ d'un processus $(B_t)_t$.

Définition 2.1.3 *On dit que le processus $(X_t^n, t \geq 0)$ converge en loi au sens des marginales de dimension finie vers le processus $(B_t, t \geq 0)$.*

Le processus $(B_t, t \geq 0)$ est tel que les accroissements $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_p} - B_{t_{p-1}})$ sont indépendants et satisfont de plus que pour chaque $t > 0$, la variable aléatoire B_t suit la loi normale centrée de variance t . Le processus ainsi construit s'appelle le mouvement brownien et joue un rôle fondamental en probabilités.

2.1.2 Le mouvement brownien

Historiquement, le mouvement brownien est associé à l'analyse de mouvements dont la dynamique au cours du temps est si désordonnée qu'il semble difficile de la prévoir, même

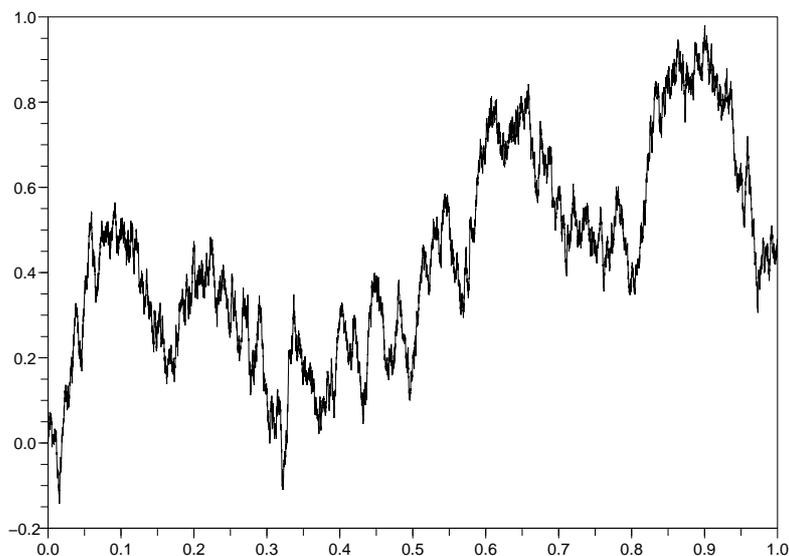


FIG. 2.3 – mouvement brownien en dimension 1

pour un temps très court. C'est Robert Brown, botaniste anglais, qui décrit en 1827 le mouvement erratique de fines particules organiques (particules de pollen) en suspension dans un gaz ou un fluide. Ce mouvement très irrégulier ne semblait pas admettre de tangente ; on ne pouvait donc pas parler de sa vitesse, ni a fortiori lui appliquer les lois de la mécanique. Le mouvement brownien fut introduit par Bachelier en 1900 pour modéliser la dynamique du prix des actions à la Bourse. Il fut redécouvert par Einstein en 1905 et est devenu l'un des outils majeurs en physique et dans la modélisation probabiliste.

Définition 2.1.4 *On appelle **processus aléatoire** $X = (X_t)_{t \geq 0}$ une famille de variables aléatoires indexée par \mathbb{R}_+ , toutes ces variables étant définies sur le même espace de probabilité. On suppose ici que chaque X_t est à valeurs réelles.*

On peut également voir ce processus comme une variable aléatoire X définie sur cet espace de probabilité et à valeurs dans l'ensemble des fonctions $t \mapsto x_t$ de \mathbb{R}_+ dans \mathbb{R} .

La théorie moderne des probabilités repose sur les résultats fondamentaux de Kolmogorov qui permettent en particulier de construire sur cet espace de fonctions une tribu qui rend l'application X mesurable (et permet donc de parler de variable aléatoire). Kolmogorov montre également que la loi de cette variable est caractérisée par ses lois marginales de dimension finie, définies comme étant les lois des k -uplets $(X_{t_1}, \dots, X_{t_k})$, pour tous temps t_1, \dots, t_k . C'est pour cela que la convergence au sens des marginales de dimension finie étudiée au paragraphe précédent permet d'en caractériser la limite.

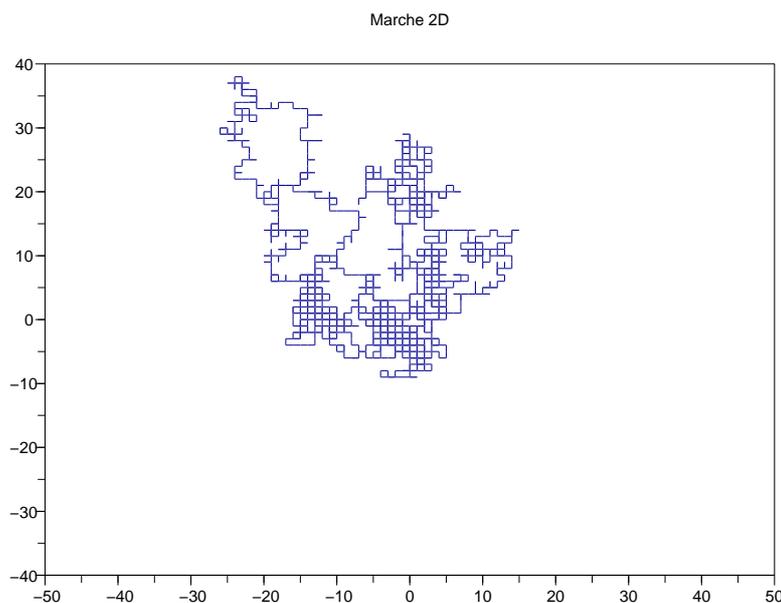


FIG. 2.4 – marche aléatoire en dimension 2

Définition 2.1.5 *Le mouvement brownien est un processus à accroissements indépendants et stationnaires, presque-sûrement à trajectoires continues, et dont la marginale au temps t a une loi normale centrée de variance t . Plus précisément, pour tous $s, t \geq 0$, la variable $B_{t+s} - B_t$ est indépendante des variables $(B_r : r \leq t)$, $B_0 = 0$, et de plus la loi de l'accroissement $B_{t+s} - B_t$ est la loi normale $\mathcal{N}(0, s)$. Elle ne dépend donc que de s .*

Remarque 2.1.6 *Nous avons montré dans le paragraphe précédent l'existence du mouvement brownien, en le construisant par approximation.*

Remarque 2.1.7 Soit $(B_t, t \geq 0)$ un mouvement brownien.

- $B'_t = -B_t$ définit un nouveau mouvement brownien.
- Pour tout $c > 0$, $B_t^{(c)} = \frac{1}{c}B_{c^2t}$ définit un nouveau mouvement brownien. Cette propriété s'appelle auto-similarité. Elle justifie l'aspect fractal des trajectoires du mouvement brownien.

Remarque 2.1.8 En fait on n'a pas besoin d'inclure la continuité des trajectoires dans la définition du mouvement brownien. On peut montrer que si un processus vérifie toutes les propriétés de la définition 2.1.5 sauf la continuité, alors il admet une modification à trajectoires continues, au sens où il existe une fonction aléatoire continue W telle que $\mathbb{P}(\text{pour chaque } t, B_t = W_t) = 1$.

Pour montrer ce résultat, on applique le critère suivant, appelé critère de Kolmogorov.

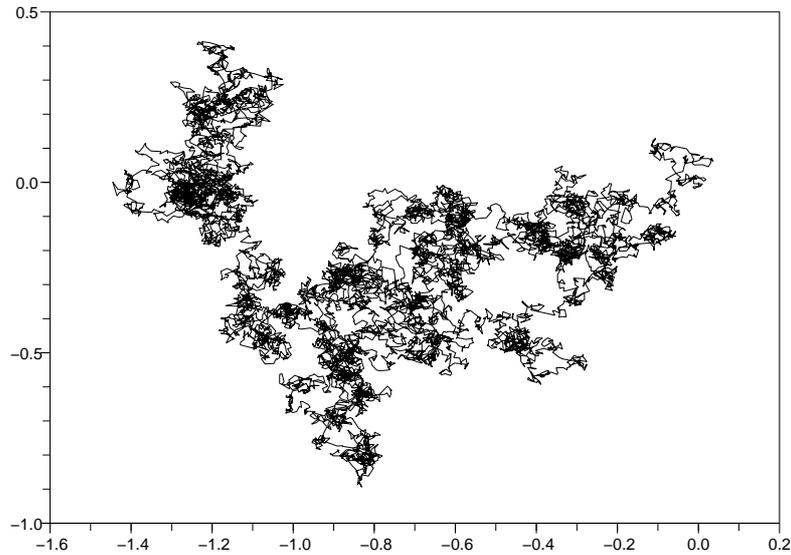


FIG. 2.5 – mouvement brownien en dimension 2

Théorème 2.1.9 Critère de Kolmogorov. (Voir Karatzas-Shreve, p.53). Si un processus X est tel qu'il existe trois constantes strictement positives α, β, C avec, pour tous t et h

$$\mathbb{E}(|X_{t+h} - X_t|^\alpha) \leq Ch^{1+\beta}, \quad (2.14)$$

alors X admet une modification presque-sûrement à trajectoires continues.

Dans le cas du mouvement brownien, la variable aléatoire $B_{t+h} - B_t$ est gaussienne centrée de variance h , donc

$$\mathbb{E}(|B_{t+h} - B_t|^4) = 3h^2.$$

En affinant le critère de Kolmogorov, on peut également montrer que les trajectoires du mouvement brownien sont höldériennes d'ordre α , pour $\alpha < 1/2$, c'est à dire que presque sûrement, $|W_{t+h} - W_t| \leq C|h|^\alpha$ pour une constante C . En dehors de ces résultats de continuité, les propriétés de régularité du mouvement brownien sont très mauvaises. Pour le montrer nous allons introduire la "variation quadratique approximée" de B au niveau n comme étant le processus

$$V(B, n)_t = \sum_{i=1}^{[nt]} (B_{i/n} - B_{(i-1)/n})^2, \quad (2.15)$$

où $[nt]$ est la partie entière de nt . On montre alors la

Proposition 2.1.10 *Pour tout t , $V(B, n)_t$ converge vers t dans \mathbb{L}^2 quand n tend vers l'infini. On appelle **variation quadratique du mouvement brownien** cette limite et on la note $\langle B, B \rangle_t = t$.*

Remarquons que cette variation quadratique est déterministe. C'est en particulier ce qui donne un caractère intrinsèque au mouvement brownien.

Preuve. Remarquons que

$$\mathbb{E} \left((V(B, n)_t - t)^2 \right) \leq 2 \left(\mathbb{E} \left(V(B, n)_t - \frac{[nt]}{n} \right)^2 + \left(\frac{[nt]}{n} - t \right)^2 \right). \quad (2.16)$$

Chaque accroissement $B_{i/n} - B_{(i-1)/n}$ a une loi normale $\mathcal{N}(0, 1/n)$, et donc $\mathbb{E}(V(B, n)_t) = \frac{[nt]}{n}$. On a

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{i=1}^{[nt]} (B_{i/n} - B_{(i-1)/n})^2 - \frac{[nt]}{n} \right)^2 \right) &= \sum_{i=1}^{[nt]} \mathbb{E} \left(\left((B_{i/n} - B_{(i-1)/n})^2 - \frac{1}{n} \right)^2 \right) \\ &= \sum_{i=1}^{[nt]} \left(\mathbb{E} \left((B_{i/n} - B_{(i-1)/n})^4 \right) - \left(\frac{1}{n} \right)^2 \right) \\ &= \frac{2}{n^2} \times [nt] \sim \frac{2t}{n} \rightarrow 0. \end{aligned}$$

On a utilisé dans ces calculs l'indépendance des variables aléatoires $(B_{i/n} - B_{(i-1)/n})^2 - \frac{1}{n}$, le développement du carré de ces variables, et le fait que $\mathbb{E} \left((B_{i/n} - B_{(i-1)/n})^2 \right) = \frac{1}{n}$, et $\mathbb{E} \left((B_{i/n} - B_{(i-1)/n})^4 \right) = \frac{3}{n^2}$.

Nous en déduisons que la variation quadratique du mouvement brownien sur $[0, t]$ est t . \square

Cela montre en particulier la

Proposition 2.1.11 *Les trajectoires $t \mapsto B_t$ sont presque sûrement nulle part dérivables.*

Pour cette raison, on ne pourra pas définir simplement l'intégrale $\int f(s) dW_s(\omega)$. Cela va justifier ultérieurement la construction de l'intégrale stochastique.

Preuve. Par le résultat précédent, en dehors d'un ensemble de probabilité nulle, on sait que pour toute paire de rationnels p, q il existe une subdivision $(t_i)_i$ de pas $\frac{1}{n}$ de $[p, q]$ telle que

$$\lim_n \sum_i (B_{t_i}(\omega) - B_{t_{i-1}}(\omega))^2 = q - p.$$

Si $|B_t(\omega) - B_s(\omega)| \leq k|t - s|$ pour $p \leq s < t \leq q$, alors on aurait

$$\sum_i (B_{t_i}(\omega) - B_{t_{i-1}}(\omega))^2 \leq k^2(q - p) \sup_i |t_i - t_{i-1}| = \frac{k^2(q - p)}{n}.$$

Quand $n \rightarrow \infty$, le fait que le terme de gauche tende vers une limite finie non nulle entraînerait une contradiction. Donc les trajectoires sont presque partout dérivables (On a utilisé la densité de \mathbb{Q} dans \mathbb{R}). \square

On peut par ailleurs décrire le comportement en temps long du mouvement brownien. Il est donné par la “loi du logarithme itéré” suivante, que l’on énonce sans démonstration.

Proposition 2.1.12 (*Karatzas-Shreve, Théorème 9.23 p.111*)

$$P \left(\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = 1 \quad \text{et} \quad \liminf_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = -1 \right) = 1. \quad (2.17)$$

2.1.3 La propriété de Markov du mouvement brownien

Définissons tout d’abord la notion de filtration, qui modélise l’évolution de l’information au cours du temps.

Soit $X = (X_t)_{t \geq 0}$ un processus sur un espace de probabilité (Ω, \mathcal{A}, P) . Nous notons \mathcal{F}_t la tribu engendrée par les variables aléatoires X_s pour $s \leq t$, c’est-à-dire la plus petite tribu rendant toutes ces variables mesurables. Nous avons $\mathcal{F}_t \subset \mathcal{A}$, et également $\mathcal{F}_s \subset \mathcal{F}_t$ si $s \leq t$.

Définition 2.1.13 *La famille croissante $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ est appelée la **filtration engendrée par le processus X** , et est aussi notée $\mathbb{F}^X = (\mathcal{F}_t^X)_{t \geq 0}$.*

La propriété de Markov

Considérons un mouvement brownien B sur (Ω, \mathcal{A}, P) , et la filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ qu’il engendre. Puisqu’il est à accroissements indépendants, la variable $Y := B_{t+s} - B_t$ est indépendante de la tribu \mathcal{F}_t et a pour loi la loi $\mathcal{N}(0, s)$. Ainsi pour chaque fonction f borélienne bornée sur \mathbb{R} ,

$$\mathbb{E}(f(B_{t+s}) | \mathcal{F}_t) = \mathbb{E}(f(B_t + Y) | \mathcal{F}_t) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi s}} e^{-x^2/2s} f(B_t + x) dx. \quad (2.18)$$

Cette formule montre que conditionnellement à \mathcal{F}_t , la loi de B_{t+s} ne dépend pas de tout le passé (i.e. de toutes les variables B_r pour $r \leq t$), mais seulement de la valeur “présente” B_t du processus. On dit que le mouvement brownien est un **processus de Markov**.

Définition 2.1.14 *Un processus $(X_t)_t$ est un processus de Markov si, étant donnée la filtration $(\mathcal{F}_t^X)_t$ engendrée par le processus, celui-ci vérifie la propriété de Markov, à savoir que pour tous $s, t \geq 0$ et pour toute fonction f borélienne bornée sur \mathbb{R} ,*

$$\mathbb{E}(f(X_{t+s})|\mathcal{F}_t^X) = \mathbb{E}(f(X_{t+s})|X_t) \quad (2.19)$$

Dans le cas du mouvement brownien, les variables B_r pour $r \geq t$ d'une part, et les variables B_r pour $r \leq t$ d'autre part, sont indépendantes, conditionnellement à la valeur de B_t . De plus, la loi de B_{t+s} sachant \mathcal{F}_t dépend bien sûr de s , mais pas de t . On dit que le mouvement brownien est un processus de Markov **homogène en temps**.

Définition 2.1.15 *Un processus $(X_t)_t$ est un processus de Markov homogène en temps si pour tous $s, t \geq 0$ et pour toute fonction f borélienne bornée sur \mathbb{R} , il existe une fonction borélienne bornée h_s telle que*

$$\mathbb{E}(f(X_{t+s})|\mathcal{F}_t^X) = h_s(X_t). \quad (2.20)$$

On peut montrer facilement que le passé et le futur d'un processus de Markov sont indépendants conditionnellement au présent.

Si $(X_t)_t$ est un processus de Markov homogène, son évolution est décrite par une famille d'opérateurs $(P_t)_t$ sur les fonctions mesurables bornées, définie pour $\phi \in \mathbb{L}^\infty$ par

$$P_t\phi(x) = \mathbb{E}(\phi(X_t)|X_0 = x).$$

Par la propriété de Markov, la famille $(P_t)_t$ vérifie la propriété de semi-groupe : $\forall \phi \in \mathbb{L}^\infty$,

$$P_{t+s}\phi = P_t(P_s\phi), \quad P_0\phi = Id.$$

Dans le cas du mouvement brownien, le semi-groupe est donné par

$$P_t\phi(x) = \mathbb{E}(\phi(B_t)|B_0 = x) = \int_{\mathbb{R}^d} \phi(y) \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) dy, \quad (2.21)$$

Notons que

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \quad (2.22)$$

est la densité au temps t du mouvement brownien issu de x au temps 0.

En utilisant la forme explicite (2.21), on peut alors montrer que si ϕ est continue, la fonction $u(t, x) = P_t\phi(x) = \int p(t, x, y)\phi(y)dy$ est l'unique solution régulière de l'équation aux dérivées partielles (2.25)

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} \quad (2.23)$$

avec la condition initiale $u(0, x) = \phi(x), x \in \mathbb{R}$.

Cette équation s'appelle **l'équation de la chaleur**.

On peut généraliser la définition de mouvement brownien au cas multi-dimensionnel.

Définition 2.1.16 *Un mouvement brownien d -dimensionnel est un d -uplet $B = (B^i)_{1 \leq i \leq d}$ de d mouvements browniens à valeurs réelles $B^i = (B_t^i)_{t \geq 0}$, qui sont **indépendants** entre eux.*

Ce processus est encore un processus de Markov homogène (et même un processus à accroissements indépendants). Son semi-groupe est alors défini par

$$P_t \phi(x) = \int_{\mathbb{R}^d} \phi(y) \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{\|y-x\|^2}{2t}\right) dy, \quad (2.24)$$

où x et y appartiennent à \mathbb{R}^d , $\|\cdot\|$ désigne la norme euclidienne sur \mathbb{R}^d , et dy la mesure de Lebesgue sur \mathbb{R}^d . Comme en dimension 1, il donne encore la solution explicite de l'équation de la chaleur en dimension d :

$$\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u \quad (2.25)$$

avec la condition initiale $u(0, x) = \phi(x), x \in \mathbb{R}^d$

2.2 Martingales et temps d'arrêt

En utilisant (2.18) et (2.21) on obtient facilement les propriétés

$$\mathbb{E}(B_t | \mathcal{F}_s) = B_s, \quad \forall s \leq t; \quad (2.2.1)$$

$$\mathbb{E}(B_t^2 - t | \mathcal{F}_s) = B_s^2 - s, \quad \forall s \leq t. \quad (2.2.2)$$

2.2.1 Martingales

Nous allons nous focaliser sur la propriété (2.2.1) satisfaite par le mouvement brownien. Elle signifie que le mouvement brownien modélise des fluctuations aléatoires qui sont constantes en espérance, et ce quelque soit l'information que l'on a à un temps précis. La propriété (2.2.1) va permettre de définir une classe de processus fondamentale en probabilité : les martingales.

Notons comme précédemment par $(\mathcal{F}_t)_t$ la filtration engendrée par le mouvement brownien. Tous les processus (X_t) que l'on va considérer sont supposés **adaptés** à la filtration $(\mathcal{F}_t)_{t \geq 0}$, ce qui veut dire que pour tout t , X_t est \mathcal{F}_t -mesurable : l'information concernant X_t est obtenue à partir de l'information donnée par la trajectoire brownienne jusqu'au temps t .

Définition 2.2.1 *Un processus à valeurs réelles $M = (M_t)_{t \geq 0}$ est une **martingale** si chaque variable M_t est intégrable, et si*

$$s \leq t \quad \Rightarrow \quad \mathbb{E}(M_t | \mathcal{F}_s) = M_s. \quad (2.2.3)$$

En particulier, l'espérance $\mathbb{E}(M_t)$ d'une martingale est une fonction constante.

Cette notion généralise la notion de martingale à temps discret (cf. MAP 432). Toutes les martingales que l'on verra dans le cours auront des trajectoires qui seront au moins continues à droite et avec des limites à gauche. (Pour le mouvement brownien, elles sont continues).

Exemples

Les propriétés (2.2.1) et (2.2.2) entraînent que

1. Le processus B est une martingale.
2. Le processus $M_t = B_t^2 - t$ est une martingale.

On peut également montrer que

3. Soit $\lambda \in \mathbb{R}$. Le processus $M_t^\lambda = \exp(\lambda B_t - \frac{\lambda^2}{2}t)$ est une martingale. En effet on peut écrire

$$\mathbb{E}(M_{t+s}^\lambda | \mathcal{F}_s) = M_t^\lambda \mathbb{E}\left(\exp(\lambda(B_{t+s} - B_t) - \frac{\lambda^2}{2}s)\right)$$

puisque $B_{t+s} - B_t$ est indépendant of \mathcal{F}_s . On utilise alors le fait que si U est une variable $\mathcal{N}(0, \sigma^2)$, alors $\mathbb{E}(\exp(\lambda U)) = \exp\left(\frac{\lambda^2 \sigma^2}{2}\right)$.

Ainsi, $\mathbb{E}\left(\exp(\lambda(B_{t+s} - B_t) - \frac{\lambda^2}{2}s)\right) = 1$ et $\mathbb{E}(M_{t+s}^\lambda | \mathcal{F}_s) = M_t^\lambda$.

Le comportement d'une martingale à l'infini

Un gros intérêt de la notion de martingale est le théorème suivant, qui montre que l'on connaît le comportement qualitatif d'une martingale en temps long.

Théorème 2.2.2 (cf. Karatzas-Shreve [13] p. 17) *Si M est une martingale satisfaisant*

$$\sup_{t \in \mathbb{R}_+} \mathbb{E}(|M_t|) < +\infty, \quad (2.2.4)$$

alors avec probabilité 1

$$M_t \rightarrow_{t \rightarrow +\infty} M_\infty \in]-\infty, +\infty[, \quad \text{et} \quad \mathbb{E}(|M_\infty|) < \infty.$$

Observons que cette propriété ne s'applique pas à l'exemple 1 ci-dessus, pour lequel $\mathbb{E}(|B_t|) = \sqrt{2t/\pi}$, ni pour l'exemple 2, pour lequel $\mathbb{E}(|M_t|) = ct$ pour une constante $c > 0$, tandis qu'il s'applique pour l'exemple 3, puisque $\mathbb{E}(|M_t^\lambda|) = \mathbb{E}(M_t^\lambda) = 1$.

En fait, dans ce dernier cas, on peut montrer, en utilisant la loi du logarithme itéré (énoncée en Proposition 2.1.12), que $M_\infty^\lambda = 0$ si $\lambda \neq 0$, et que par ailleurs, $M_\infty^\lambda = 1$ si $\lambda = 0$, puisque $M_t^\lambda = 1$ pour tout t . Ainsi, l'égalité (2.2.3) est clairement fautive quand $s = \infty$, sauf dans le cas où $\lambda = 0$.

Une importante question, mais trop difficile dans le cadre de ce cours, est alors de savoir quand l'égalité (2.2.3) reste vraie pour $s = \infty$. Comme nous l'avons vu grâce à l'exemple 3, les hypothèses doivent être strictement plus fortes que l'hypothèse $\sup_{t \in \mathbb{R}_+} \mathbb{E}(|M_t|) < +\infty$. Nous nous contenterons de la réponse partielle suivante.

Théorème 2.2.3 ([13] Problème 3.19, p.18) *Si la martingale M est bornée :*

$$\forall \omega, \forall t, |M_t(\omega)| \leq K, \quad (2.2.5)$$

alors

$$M_t = E(M_\infty | \mathcal{F}_t).$$

2.2.2 Temps d'arrêt

Les temps d'arrêt jouent un rôle très important en théorie des probabilités. Nous généralisons ici la notion introduite en temps discret.

Définition 2.2.4 *Une variable aléatoire $T \in [0, \infty]$ est un **temps d'arrêt** si l'événement $\{T \leq t\} \in \mathcal{F}_t$ pour tout $t \geq 0$.*

Un temps d'arrêt est donc un temps aléatoire, tel que sur chaque ensemble $\{\omega : T(\omega) \leq t\}$, $T(\omega)$ dépend seulement de l'information donnée par le mouvement brownien avant le temps t .

Un exemple trivial de temps d'arrêt est donné par $T(\omega) = t$ pour tout ω .

En dehors des temps constants, l'exemple fondamental de temps d'arrêt est le **temps d'atteinte** d'un ensemble borélien A par un processus X . On définit plus précisément

$$T_A = \inf\{t \geq 0; X_t \in A\}$$

(avec la convention que l'infimum de l'ensemble vide vaut $+\infty$). Même si la preuve prouvant que T_A est un temps d'arrêt est délicate, l'idée intuitive est simple : on n'a besoin que de la connaissance de la trajectoire de $X(\omega)$ jusqu'au temps t pour savoir si $T_A(\omega) \leq t$.

Ce raisonnement permet de comprendre aussi pourquoi en revanche, le dernier temps avant un temps fixé s où X visite A , défini par $S = \sup\{t \geq 0, X_t \in A\}$ (où le supremum de l'ensemble vide est égal à 0) n'est pas un temps d'arrêt. En effet, la valeur de $T(\omega)$ dépend de toute la trajectoire $X(\omega)$.

Le théorème d'arrêt

Théorème 2.2.5 *Soit M une martingale. La propriété (2.2.3) peut facilement être étendue aux temps d'arrêt bornés. En particulier, si S et T sont deux temps d'arrêt bornés, alors*

$$\mathbb{E}(M_T) = \mathbb{E}(M_S). \quad (2.2.6)$$

*En particulier, si T est un temps d'arrêt **borné**, on a*

$$\mathbb{E}(M_T) = \mathbb{E}(M_0). \quad (2.2.7)$$

Observons que (2.2.7) est fautive si T n'est pas borné. Par exemple si $M = B$ est un mouvement brownien et si $T = \inf\{t : M_t = 1\}$, alors $\mathbb{E}(M_0) = 0 < \mathbb{E}(M_T) = 1$. Dans ce cas, le temps aléatoire T est presque sûrement fini, mais n'est pas borné et a même une espérance infinie (voir la section 5.3.2 ci-dessous).

En revanche, dans le cas d'une martingale bornée (c-à-d vérifiant (2.2.5)), tout se passe bien.

Théorème 2.2.6 *Si M est une martingale bornée, alors (2.2.6) et (2.2.7) ont lieu pour tous temps d'arrêt S et T .*

Remarquons que l'on peut considérer des temps d'arrêt qui peuvent prendre la valeur infinie, pourvu que l'on pose $M_T = M_\infty$ sur l'ensemble $\{T = \infty\}$.

Considérons une martingale M et un temps d'arrêt T . Définissons la martingale stoppée au temps T par $M_t^T = M_{\min(t,T)}$. On peut facilement déduire du Théorème 2.2.5 qu'alors

$$M^T \text{ est une martingale.} \quad (2.2.8)$$

Définition 2.2.7 *Si M est une martingale et T un temps d'arrêt, on appelle **martingale arrêtée au temps T** la martingale M^T .*

2.2.3 Applications au mouvement brownien

A - La propriété de Markov forte

Définition 2.2.8 *Soit X un processus de Markov homogène, de semi-groupe de transition $(P_t)_{t \geq 0}$. On dit que X est un processus **fortement markovien** si pour chaque temps d'arrêt T , conditionnellement à l'information sur la trajectoire jusqu'au temps T et sur l'ensemble $\{T < \infty\}$, la loi de X_{T+s} dépend seulement de X_T .*

La plupart du temps, les processus de Markov homogènes sont fortement markoviens.

On peut montrer le

Théorème 2.2.9 *(voir [13], Section 2.6) Le mouvement brownien est fortement markovien. De plus, si T est un temps d'arrêt, alors sur l'ensemble $\{T < \infty\}$, le processus $B'_s = B_{T+s} - B_T$ est encore un mouvement brownien, indépendant de \mathcal{F}_T .*

B - Temps d'atteinte pour le mouvement brownien**1) Calcul de la loi du temps d'atteinte d'un point.**

Soit B un mouvement brownien, et soit pour $a > 0$ le temps d'arrêt

$$T_a = \inf\{t : B_t = a\}.$$

Considérons la martingale $M_t^\lambda = \exp(\lambda B_t - \frac{\lambda^2}{2}t)$ pour un $\lambda > 0$ arbitraire, et définissons la martingale $N = (M^\lambda)^{T_a}$ arrêtée au temps T_a , et définie par

$$N_t = (M^\lambda)_t^{T_a} = (M^\lambda)_{T_a \wedge t}.$$

On a donc $0 < N_t \leq e^{a\lambda}$, ce qui entraîne que N est bornée. On peut donc appliquer (2.2.7) à la martingale N et au temps d'arrêt T_a , ce qui donne $\mathbb{E}(N_{T_a}) = \mathbb{E}(N_0) = 1$. Puisque $N_{T_a} = \exp(\lambda a - \frac{\lambda^2}{2}T_a)$, et si l'on pose $\vartheta = \lambda^2/2$, on obtient

$$\mathbb{E}(e^{-\vartheta T_a}) = e^{-a\sqrt{2\vartheta}}. \quad (2.2.9)$$

On obtient ainsi la “transformée de Laplace” de la variable T_a . Cette transformée de Laplace peut être inversée, et cela montre que T_a admet une densité sur \mathbb{R}_+ donnée par

$$f_a(y) = \frac{a}{\sqrt{2\pi y^3}} e^{-a^2/2y}, \quad y \geq 0. \quad (2.2.10)$$

La loi de T_a s'appelle **une loi “stable” d'indice 1/2**. En particulier on en déduit que

$$\mathbb{P}(T_a < \infty) = 1 \quad \text{et} \quad \mathbb{E}(T_a) = \infty.$$

Remarquons que ces résultats peuvent être obtenus simplement à partir de (2.2.9), le premier en faisant tendre ϑ vers 0 et le deuxième en dérivant par rapport à ϑ et en faisant tendre ϑ vers 0.

Des formules similaires existent avec $-a$ au lieu de a , par symétrie du mouvement brownien. Le processus $-B$ est encore un mouvement brownien, et le temps d'atteinte de a par $-B$ est égal au temps d'atteinte de $-a$ par B .

Considérons maintenant $a > 0$ et $b > 0$, et soit $T = \min(T_a, T_{-b})$. La martingale arrêtée $M_t = B_{\min(T, t)}$ est bornée, donc on peut appliquer (2.2.7) au temps T . Cela entraîne que

$$0 = \mathbb{E}(M_0) = \mathbb{E}(M_T) = a\mathbb{P}(T = T_a) - b\mathbb{P}(T = T_{-b}).$$

Mais $\{T = T_a\} = \{T_a < T_{-b}\}$ et $\{T = T_{-b}\} = \{T_{-b} < T_a\}$, de telle sorte que finalement on obtient

$$\mathbb{P}(T_a < T_{-b}) = \frac{b}{a+b}, \quad \mathbb{P}(T_{-b} < T_a) = \frac{a}{a+b}. \quad (2.2.11)$$

2) Le principe de réflexion.

Introduisons le processus continu croissant $S_t = \sup_{s \leq t} B_s$. Ce processus et les temps d'arrêt T_a sont inverses l'un de l'autre dans le sens où $\{T_a \leq t\} = \{S_t \geq a\}$, et donc

$$T_a = \inf\{t; S_t \geq a\} \text{ et } S_t = \inf\{a, T_a \geq t\}.$$

La propriété de Markov forte permet de montrer le "principe de réflexion".

Proposition 2.2.10 *Pour $a > 0$,*

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(T_a \leq t) = 2\mathbb{P}(B_t \geq a) = \mathbb{P}(|B_t| \geq a) = 2(1 - \Phi(\frac{a}{\sqrt{t}})), \quad (2.2.12)$$

où Φ désigne la fonction de répartition de la loi $\mathcal{N}(0, 1)$, définie par $\Phi(t) = P(U \leq t)$ si U est une variable normale centrée réduite.

Le nom de ce résultat vient de l'argument heuristique suivant. Parmi les trajectoires qui atteignent a avant le temps t , "une moitié" sera supérieure à a au temps t . En effet, en utilisant la trajectoire symétrique par rapport à la droite " $y = a$ " de la trajectoire du mouvement brownien entre les temps T_a et t , on peut montrer l'existence d'une correspondance bijective et préservant les probabilités entre les trajectoires supérieures à a au temps t , et celles qui sont inférieures à a au temps t .

Preuve.

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(T_a \leq t, B_t \geq a) + \mathbb{P}(T_a \leq t, B_t < a).$$

Mais $\mathbb{P}(T_a \leq t, B_t \geq a) = \mathbb{P}(B_t \geq a)$. Par ailleurs,

$$\begin{aligned} \mathbb{P}(T_a \leq t, B_t < a) &= \mathbb{E}(T_a \leq t, \mathbb{E}(B_t < a | \mathcal{F}_{T_a})) \\ &= \mathbb{P}(T_a \leq t) \mathbb{E}(B_t < a | B_0 = a) = \frac{1}{2} \mathbb{P}(T_a \leq t), \end{aligned}$$

par la propriété de Markov forte et par symétrie du mouvement brownien. On en déduit le résultat. \square

3) Récurrence du mouvement brownien.

Finalement, puisque $\mathbb{P}(T_a < \infty) = 1$, et de nouveau à cause de la propriété de Markov forte, on vérifie facilement ce qu'on appelle la **récurrence** du mouvement brownien.

Proposition 2.2.11 *Presque sûrement, pour tout $t \geq 0$, le processus B visite infiniment souvent chaque nombre réel a après le temps t .*

2.2.4 Applications aux temps d'atteinte de barrières

Comme dans le cas des marches aléatoires nous allons nous intéresser à des barrières absorbantes ou réfléchissantes.

Barrière absorbante en 0

Soit $x > 0$. Considérons un organisme planctonique issu de $x > 0$ (la position 0 est le fond de la mer) et diffusant verticalement suivant un mouvement brownien. On considère que la profondeur est suffisamment grande par rapport à la taille de la cellule pour que son déplacement soit considéré comme possible sur tout \mathbb{R}_+ . Le fond de la mer est recouvert de moules (prédatrices de plancton), et la particule est absorbée dès qu'elle touche le fond. Cherchons le temps moyen de capture d'un individu issu d'une profondeur $x > 0$. Avant d'être absorbé, l'individu suit un mouvement brownien issu de x , soit $X_t = x + B_t$. On note encore $T_0 = \inf\{t : X_t = 0\}$, et l'on souhaite calculer le temps moyen d'absorption

$$m(x) = \mathbb{E}(T_0),$$

sachant que $X_0 = x$. En adaptant la preuve précédente, nous obtenons immédiatement que

$$m(x) = +\infty.$$

Supposons maintenant que la surface de la mer soit à la hauteur $a > 0$ et que le phytoplancton, quand il arrive à la surface y trouve une nappe de pétrole et y soit absorbé. Il va alors se déplacer suivant un mouvement brownien issu de $x \in [0, a]$, absorbé en 0 et en a . Ainsi, dans ce cas,

$$m(x) = \mathbb{E}(T_0 \wedge T_a).$$

Notons $T = T_0 \wedge T_a$. Le processus $X_t^T = X_{t \wedge T}$ est une martingale issue de x , bornée par 0 et a . C'est donc une martingale bornée. Nous pouvons alors appliquer un raisonnement analogue à ceux du paragraphe précédent. Nous obtenons que

$$\mathbb{E}(X_T) = a\mathbb{P}(X_T = a) = a\mathbb{P}(T_a < T_0) = \mathbb{E}(X_0) = x.$$

Ainsi,

$$\mathbb{P}(T_a < T_0) = \frac{x}{a} \quad ; \quad \mathbb{P}(T_0 < T_a) = 1 - \frac{x}{a},$$

puisque le processus est récurrent et que donc $\mathbb{P}(T < \infty) = 1$. On peut alors utiliser la martingale $X_t^2 - t$ pour trouver l'espérance de T . Nous savons que $X_t^2 - t$ est un martingale. Toutefois elle n'est pas bornée pour $t \leq T$. En revanche nous allons pouvoir appliquer le théorème d'arrêt au temps d'arrêt $T \wedge k$, pour k un entier quelconque fixé. Le processus $X_{t \wedge T \wedge k}^2 - (t \wedge T \wedge k)$ est une martingale bornée. Nous obtenons alors :

$$\begin{aligned} \mathbb{E}((X_{t \wedge k})^2) &= x^2 + \mathbb{E}(T \wedge k) \\ &= a^2 \mathbb{P}(T_a < T_0, T_a < k) + \mathbb{E}((X_k)^2 \mathbf{1}_{k < T}). \end{aligned}$$

Mais nous savons que $P(T < \infty) = 1$, d'où

$$\mathbb{E}((X_k)^2 \mathbf{1}_{k < T}) \leq a^2 \mathbb{P}(k < T),$$

et cette dernière quantité tend vers 0 quand k tend vers l'infini. Ainsi, quand k tend vers l'infini, nous obtenons que

$$a^2 \mathbb{P}(T_a < T_0) = x^2 + \mathbb{E}(T)$$

d'où finalement

$$\mathbb{E}(T) = x(a - x).$$

On peut également facilement montrer que

$$\mathbb{E}(T_a | T_a < T_0) = \frac{a(a - x)}{2} \quad ; \quad \mathbb{E}(T_0 | T_0 < T_a) = \frac{ax}{2}.$$

Supposons maintenant que la surface de la mer est à la hauteur a et que la cellule de plancton se réfléchit naturellement à la surface. Elle va alors se déplacer suivant un mouvement brownien issu de x , absorbé en 0 et réfléchi en a . Mais si le mouvement brownien atteint a avant 0, alors après réflexion il va se comporter comme un nouveau mouvement brownien issu de a et indépendant du précédent allant dans le sens inverse (propriété de Markov forte). Pour ce deuxième mouvement brownien, la probabilité d'atteindre 0 en partant de a va alors être la même que la probabilité d'atteindre $2a$ pour le premier (en partant de a). Nous allons adapter les calculs du cas précédent en remplaçant a par $2a$. Ainsi, dans ce cas,

$$m(x) = \mathbb{E}(T_0 \wedge T_{2a} | X_0 = x),$$

où $0 < x < a$. On aura

$$\mathbb{P}(T_{2a} < T_0) = \frac{x}{2a} \quad ; \quad \mathbb{P}(T_0 < T_{2a}) = 1 - \frac{x}{2a},$$

et

$$m(x) = \mathbb{E}(T) = x(2a - x).$$

2.3 Intégrales stochastiques et équations différentielles stochastiques

Nous avons vu à la Remarque 2.1.2 que certaines normalisations de marches aléatoires pouvaient conduire à des processus de type $t \rightarrow \mu t + \sigma B_t$, somme d'un processus déterministe et d'un processus aléatoire. Si $\mu \neq 0$, le premier terme, appelé terme de dérive, peut être vu comme un terme de tendance centrale et le terme brownien, appelé terme de diffusion comme une perturbation aléatoire. Nous allons ici généraliser ce type de processus.

Usuellement, quand les fluctuations aléatoires sont inexistantes ou extrêmement petites, la description du mouvement d'une particule est décrite par la solution d'une équation différentielle ordinaire, de type

$$dx_t = b(x_t)dt.$$

Supposons que le mouvement soit perturbé par une composante aléatoire. Comme nous l'avons vu ci-dessus, nous pouvons modéliser l'alea des fluctuations d'une particule évoluant comme une marche aléatoire sur le temps infinitésimal dt par un mouvement brownien. Si nous supposons de plus que l'écart-type des fluctuations dépend de la position de la particule au temps t , le mouvement aléatoire de la particule sera une fonction aléatoire $t \rightarrow X_t$, solution d'une équation de la forme

$$dX_t = b(x_t)dt + \sigma(X_t)dB_t. \quad (2.3.1)$$

Ce type d'équation est appelée **équation différentielle stochastique**. On pourrait croire que c'est une généralisation d'une équation différentielle ordinaire mais en fait, elle n'a pas le même sens, puisque nous avons vu que B_t n'est nulle part dérivable en t .

Si nous supposons que la condition initiale est une variable aléatoire X_0 , une solution de (2.3.1) est par définition la solution de l'équation intégrale

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s. \quad (2.3.2)$$

Un problème majeur est de donner un sens à l'intégrale $\int_0^t \sigma(X_s)dB_s$, qui va être une intégrale de nature différente d'une intégrale de Riemann. Cette intégrale, appelée **intégrale stochastique**, a été construite par Itô en 1944 (donc très récemment).

2.3.1 Intégrales stochastiques

Nous souhaitons donner un sens à l'intégrale $\int_0^t H_s dB_s$ quand B est un mouvement brownien et $H = (H_t)_{t \geq 0}$ un processus dont les propriétés sont à préciser.

Nous nous limiterons essentiellement à des intégrands H qui sont continus en la variable t . L'idée naturelle est d'obtenir $\int_0^t H_s dB_s$ comme limite de sommes de Riemann

$$I(H, n)_t = \sum_{i=1}^{[nt]} H_{(i-1)/n} (B_{i/n} - B_{(i-1)/n}). \quad (2.3.3)$$

La variable $B_{i/n} - B_{(i-1)/n}$ a une taille d'ordre $1/\sqrt{n}$, car elle est centrée et de variance $1/n$, et donc $H_{(i-1)/n} (B_{i/n} - B_{(i-1)/n})$ est également d'ordre $1/\sqrt{n}$. La taille de $I(H, n)_t$ devrait donc être \sqrt{n} . Mais dans ce cas, les variables $I(H, n)_t$ ne pourront en général pas converger; cela est cohérent avec le fait que les trajectoires de B sont à variation infinie.

Pourtant, une sorte de “miracle” a lieu, quand on suppose de plus que le processus H est adapté à la filtration du mouvement brownien. Par simplicité, on supposera aussi que ce processus H est borné par une constante, mais cette hypothèse peut être allégée.

Dans ce cas, la variable

$$Y(n, i) = H_{(i-1)/n}(B_{i/n} - B_{(i-1)/n})$$

satisfait $\mathbb{E}(Y(n, i) | \mathcal{F}_{(i-1)/n}) = 0$ et $\mathbb{E}(Y(n, i)^2 | \mathcal{F}_{(i-1)/n}) = H_{(i-1)/n}^2/n \leq C^2/n$, de telle sorte que $I(H, n)_t = \sum_{i=1}^{[nt]} Y(n, i)$ est centré, de variance

$$\sum_{i=1}^{[nt]} \mathbb{E}(Y(n, i)^2) \leq C^2 t.$$

Il n'est alors pas totalement déraisonnable de penser que la suite $I(H, n)_t$ converge, et effectivement on peut montrer le théorème suivant.

Théorème 2.3.1 *Soit H un processus borné, continu et adapté à la filtration du mouvement brownien B . Alors, la suite*

$$I(H, n)_t = \sum_{i=1}^{[nt]} H_{(i-1)/n}(B_{i/n} - B_{(i-1)/n})$$

converge dans \mathbb{L}^2 , quand n tend vers l'infini, vers une limite, notée

$$\int_0^t H_s dB_s$$

et appelée l'intégrale stochastique de H par rapport à B sur l'intervalle $[0, t]$.

Remarque 2.3.2 1. Si H n'est pas adapté, les sommes de Riemann ne convergent pas en général. Si on remplace $H_{(i-1)/n}(B_{i/n} - B_{(i-1)/n})$ par $H_{t(n,i)}(B_{i/n} - B_{(i-1)/n})$, avec $(i-1)/n \leq t(n,i) \leq i/n$, comme il est possible de le faire pour les approximations par les sommes de Riemann pour les intégrales usuelles, la suite associée $I(H, n)_t$ ne converge pas nécessairement, et si elle converge, la limite peut être différente de $\int_0^t H_s dB_s$.

2. La terminologie “intégrale stochastique” permet d'insister sur le fait que cette intégrale n'est pas une intégrale usuelle, prise séparément pour chaque valeur de ω , mais une limite dans \mathbb{L}^2 .

Propriétés de l'intégrale stochastique

Théorème 2.3.3 (cf. [13], Chapter 3.2) *Soit H et K des processus bornés, continus et adaptés.*

1. Pour tous réels α, β , $\int_0^t (\alpha H_s + \beta K_s) dB_s = \alpha \int_0^t H_s dB_s + \beta \int_0^t K_s dB_s$.
2. Le processus M défini par $M_t = \int_0^t H_s dB_s$ est une martingale continue de carré intégrable, nulle en 0, et donc d'espérance nulle.
3. Si de plus, $N_t = \int_0^t K_s dB_s$, on a

$$M_t N_t - \int_0^t H_s K_s ds \quad (2.3.4)$$

est une martingale continue nulle en 0. On a donc

$$\mathbb{E}(M_t N_t) = \mathbb{E} \left(\int_0^t H_s K_s ds \right). \quad (2.3.5)$$

En particulier, on a l'isométrie fondamentale, donnée par la formule suivante

$$\mathbb{E} \left(\left(\int_0^t H_s dB_s \right)^2 \right) = \mathbb{E} \left(\int_0^t H_s^2 ds \right). \quad (2.3.6)$$

2.3.2 Equations différentielles stochastiques (EDS)

Définition 2.3.4 Soit B un mouvement brownien sur un espace de probabilité, et $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ la filtration qu'il engendre. Les coefficients b et σ , de même que la condition initiale $x_0 \in \mathbb{R}$, sont donnés. Nous appelons **solution forte** de (2.3.1) tout processus X , continu, adapté à la filtration \mathcal{F} , et tel que (2.3.2) ait lieu.

Observons que l'adaptation de X est nécessaire pour que l'intégrale stochastique dans (2.3.1) ait un sens.

Cherchons des conditions suffisantes pour avoir existence et unicité d'une solution.

Pour le cas purement déterministe d'une équation différentielle ordinaire, de la forme

$$dX_t = b(X_t)dt, \quad X_0 = x_0, \quad (2.3.7)$$

où le coefficient b et la condition initiale x_0 sont donnés, un résultat classique (le théorème de Cauchy-Lipschitz) énonce que (2.3.7) admet une et une seule solution dès que b est lipschitzienne. Pour l'EDS (2.3.1), on peut prouver essentiellement le même résultat.

Théorème 2.3.5 (cf. [13], Théorème 2.9 p. 289) Si les coefficients b et σ sont lipschitziens et si la condition initiale X_0 est dans \mathbb{L}^2 , alors pour tout $T > 0$, il existe une et une seule solution forte X dans l'espace $\{X, \mathbb{E}(\sup_{t \leq T} |X_t|^2) < \infty\}$, à l'EDS

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s.$$

Ce processus sera appelé **une diffusion**.

Remarque 2.3.6 Nous avons défini une EDS en dimension 1 mais on peut généraliser cette notion en dimension d , avec $b(x)$ un vecteur de \mathbb{R}^d , $\sigma(x)$ une matrice de $\mathbb{R}^d \times \mathbb{R}^d$ et B un mouvement brownien d -dimensionnel.

Remarque 2.3.7 1) L'unicité est comprise au sens presque sûr : si X et X' sont deux solutions fortes, alors $P(X_t = X'_t \text{ pour tout } t) = 1$.

2) Le fait que l'on obtienne une solution dans un espace de type \mathbb{L}^2 n'est pas surprenant : c'est en effet dans ce cadre que l'on a développé le calcul stochastique.

Comme pour les équations différentielles ordinaires, l'hypothèse de lipschitzianité est suffisante, mais pas nécessaire, pour obtenir l'existence et l'unicité de la solution. Par exemple, on a le résultat plus fort suivant, spécifique à la dimension 1.

Théorème 2.3.8 (cf. [13] Proposition 2.13 p. 291) Si

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)|^2 \leq C|x - y|, \quad (2.3.8)$$

alors il existe une unique solution forte.

Ici, σ n'est plus lipschitzienne, mais seulement höldérienne de rapport 1/2. Nous verrons plusieurs exemples d'équations de ce type ultérieurement.

2.3.3 Propriété de Markov

Soit X la solution forte de l'EDS (2.3.2) avec coefficients lipschitziens.

Fixons par ailleurs un temps T . On peut alors également résoudre l'EDS suivante,

$$dX'_t = b(X'_t)dt + \sigma(X'_t)dB_t^T, \quad X'_0 = X_T, \quad (2.3.9)$$

où $B_t^T = B_{T+t} - B_T$. Le processus B^T est encore un mouvement brownien, indépendant du passé avant T , et donc indépendant de X_T . Les coefficients dans l'équation (2.3.9) étant lipschitziens, X' existe et est unique.

Définissons maintenant un nouveau processus X'' par $X''_t = X_t$ si $t < T$ et $X''_t = X'_{t-T}$ si $t \geq T$. Clairement, au vu de (2.3.9), X'' satisfait l'équation (2.3.2). Ainsi, par l'unicité, on obtient $X'' = X$, et $X_{T+t} = X'_t$ si $t \geq 0$.

Il s'en suit que la variable aléatoire X_{T+t} dépend seulement de X_T et de B^T , et la loi conditionnelle de X_{T+t} sachant que $X_T = x$ sera égale à la loi de X_t , quand la condition initiale est $X_0 = x$. Ainsi, on en déduit le théorème suivant.

Théorème 2.3.9 *Le processus X , solution de l'EDS (2.3.2), est un processus de Markov homogène.*

On peut reproduire l'argument précédent en prenant comme temps T un temps d'arrêt fini au lieu d'un temps fixe, et nous obtenons aussi la propriété de Markov forte pour X . On note par $(P_t)_{0 \leq t}$ son semi-groupe, c'est à dire que $P_t f(x) = \mathbb{E}(f(X_t) | X_0 = x)$. Sous certaines hypothèses, on peut montrer que la loi de X_t sachant que $X_0 = x$ admet une densité.

Théorème 2.3.10 *Si $\sigma(x) > 0$ pour tout x , alors pour tout $t > 0$, il existe une fonction $(x, y) \mapsto p_t(x, y)$ sur $\mathbb{R} \times \mathbb{R}$, telle que $P_t f(x) = \int p_t(x, y) f(y) dy$, (et ceci même si la condition initiale est déterministe).*

Cet effet de régularisation est dû au mouvement brownien.

2.3.4 Formule d'Itô

Quand $t \mapsto x(t)$ est une fonction réelle, continue et à variation finie, la formule d'intégration par parties implique que pour toute fonction f continûment dérivable, on a

$$f(x(t)) = f(x(0)) + \int_0^t f'(x(s)) dx(s). \quad (2.3.10)$$

Celle-ci devient fausse quand la fonction x est remplacée par un mouvement brownien B . En effet, si (2.3.10) était vraie, en prenant $f(x) = x^2$, on obtiendrait $B_t^2 = 2 \int_0^t B_s dB_s$ et, puisque le processus défini par l'intégrale stochastique est encore une martingale, on pourrait en déduire que B_t^2 est une martingale. Or, on a vu au paragraphe précédent que la variation quadratique de B au point t est égale à t , et donc le processus $B_t^2 - t$ est aussi une martingale. Par différence, le "processus" t devrait alors être une martingale, ce qui est évidemment faux.

Ainsi (2.3.10) est fausse pour le mouvement brownien. Pour obtenir une formule juste, on doit ajouter un terme de plus et supposer plus de régularité sur f .

Théorème 2.3.11 *(cf. [13], Théorème 3.3 p. 149). Considérons X solution de l'EDS (2.3.2). Soit f est une fonction de classe C^2 , alors on a*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) b(X_s) ds + \int_0^t f'(X_s) \sigma(X_s) dB_s + \frac{1}{2} \int_0^t f''(X_s) \sigma^2(X_s) ds, \quad (2.3.11)$$

où f'' est la seconde dérivée de f .

Cette formule est connue sous le nom de "formule d'Itô" et est extrêmement utile.

Par exemple, la formule d'Itô appliquée au mouvement brownien $X = B$ et à $f(x) = x^2$ donne

$$B_t^2 = t + 2 \int_0^t B_s dB_s, \quad (2.3.12)$$

et l'on retrouve ainsi la propriété que $B_t^2 - t$ est une martingale.

Cette formule nous permet de montrer les liens entre solution d'une EDS et solution d'une équation aux dérivées partielles, liens qui nous seront utiles pour faire des calculs explicites liés à ces diffusions.

2.3.5 Générateur - Lien avec les équations aux dérivées partielles

Soit X la solution de l'EDS (2.3.1). Pour toute fonction bornée de classe C^2 à dérivées bornées, on peut assurer que $\int_0^t f'(X_s)\sigma(X_s)dB_s$ est une martingale. Nous en déduisons alors, en prenant l'espérance de chaque terme dans (2.3.11) que

$$\mathbb{E}(f(X_t)) = \mathbb{E}(f(X_0)) + \int_0^t \mathbb{E}(f'(X_s)b(X_s) + \frac{1}{2}\sigma^2(X_s)f''(X_s)) ds.$$

On aura donc

$$P_t f(x) = f(x) + \int_0^t P_s (f'b + \frac{1}{2}\sigma^2 f'')(x) ds,$$

et on peut en particulier montrer que si f est suffisamment régulière, la fonction $t \mapsto P_t f(x)$ est dérivable en 0 pour chaque x , et l'on définit le **générateur infinitésimal** A en posant

$$Af(x) = \lim_{t \rightarrow 0, t > 0} \frac{P_t f(x) - f(x)}{t} = f'(x)b(x) + \frac{1}{2}\sigma^2(x)f''(x). \quad (2.3.13)$$

On obtient donc l'équation dite *forward* :

$$\frac{d}{dt} P_t f(x) = P_t (Af)(x).$$

De plus, si la loi de X_t issu de x a une densité $p_t(x, y)$, (c'est le cas si $\sigma(y) > 0$), nous aurons alors que

$$\begin{aligned} \int f(y)p_t(x, y)dy &= f(x) + \int_0^t \int \left(f'(y)b(y) + \frac{1}{2}\sigma^2(y)f''(y) \right) p_s(x, y)dy ds, \\ &= f(x) + \int f(y) \int_0^t \left(-\frac{\partial}{\partial y}(b(y)p_s(x, y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y)p_s(x, y)) \right) ds dy, \end{aligned}$$

pour f est à support compact, en utilisant une intégration par parties. Nous en déduisons que $(p_t(x, \cdot))$ est solution faible de l'équation aux dérivées partielles, appelée **équation de Fokker-Planck** : pour tout x ,

$$\frac{\partial p}{\partial t}(t, x, y) = -\frac{\partial}{\partial y}(b(y)p_t(x, y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y)p_t(x, y)). \quad (2.3.14)$$

Par ailleurs, nous avons vu en Section 2.1.3 que puisque X satisfait la propriété de Markov, on a

$$P_{t+s}f = P_s(P_t f).$$

Alors, en dérivant par rapport à s puis en prenant $s = 0$, on obtient une deuxième équation, dite équation *backward* :

$$\frac{d}{dt}P_t f(x) = A(P_t f)(x).$$

Ainsi, dans le cas d'une densité, nous en déduisons que

$$\begin{aligned} P_t f(x) &= f(x) + \int_0^t A(P_s f)(x) ds \\ &= f(x) + \int_0^t b(x) \frac{\partial}{\partial x} \left(\int p_t(x, y) f(y) dy \right) + \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2} \left(\int p_t(x, y) f(y) dy \right). \end{aligned}$$

On en déduit que pour tout y ,

$$\frac{\partial p}{\partial t}(t, x, y) = b(x) \frac{\partial}{\partial x} p_t(x, y) + \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2} p_t(x, y). \quad (2.3.15)$$

2.3.6 Applications aux temps d'atteinte de barrières

Considérons un intervalle $[0, a]$, $a > 0$. Reprenons dans ce cadre de diffusions les problèmes de barrières absorbantes et réfléchissantes développés en Sections 2.2.3 et 2.2.4. Supposons que $X_0 = x \in [0, a]$. Il existe des conditions explicites sur les coefficients σ et b assurant que le processus de diffusion solution de (2.3.2) est récurrent. Plaçons-nous dans ce cas. Nous savons, grâce à la propriété de récurrence de la diffusion, que si le processus part de $x \in]0, a[$, nécessairement il atteindra le bord de l'intervalle $[0, a]$ presque-sûrement.

Barrières absorbantes

Supposons que la diffusion se comporte comme précédemment à l'intérieur de l'intervalle $[0, a]$, mais qu'elle reste absorbée si elle touche le bord. Nous allons introduire T_0 qui est le temps d'atteinte de 0 et T_a qui est le temps d'atteinte de a pour ce processus de diffusion absorbé. Une question importante est de savoir quelle est la loi du temps d'atteinte d'un bord et quel bord il atteindra, et donc de quantifier, pour $x \in [0, a]$, les nombres

$$q_t(x, a) = \mathbb{P}_x(T_a > t), \quad q_\infty(x, a) = \mathbb{P}_x(T_a > T_0).$$

Remarquons que $\mathbb{P}_x(T_a > t) = \mathbb{P}_x(X_t < a) = P_t f(x)$ avec $f(y) = \mathbf{1}_{y < a}$.

Sans rentrer dans les détails techniques, on peut montrer que $q_t(x, a)$ est solution de (2.3.15), avec les conditions aux bords $q_0(x, a) = 1$ si $x \neq a$, $q_0(a, a) = 0$, $q_t(a, a) = 0$, $q_t(0, a) = 1$.

En faisant tendre t vers l'infini, on obtient que

$$q_\infty(x, a) = \mathbb{P}_x(T_0 < T_a)$$

est solution de

$$b(x)q'(x) + \frac{1}{2}\sigma^2(x)q''(x) = 0,$$

avec les conditions aux bords $q(0) = 1$; $q(a) = 0$.

Nous sommes alors amenés à trouver la solution d'une équation différentielle ordinaire.

Barrières réfléchissantes

Nous supposons maintenant que le processus part de $x \in [0, a]$, évolue comme la diffusion à l'intérieur de l'intervalle et est réfléchi instantanément quand il atteint le bord. Comme dans le cas de la marche aléatoire, on peut intuitivement penser que le processus va évoluer indéfiniment dans l'intervalle en se stabilisant quand le temps tend vers l'infini. Nous voulons en chercher la loi invariante. Celle-ci va être limite de la loi de X_t quand t tend vers l'infini. En adaptant les résultats de la section 2.3.5, nous pourrions montrer que la loi du processus réfléchi en 0 et en a est solution de l'équation de Fokker-Planck (2.3.14) avec les conditions de Neumann aux bords $\frac{\partial}{\partial y}p_t(x, 0) = \frac{\partial}{\partial y}p_t(x, a) = 0$.

Si l'on fait tendre t vers l'infini, nous obtenons que la densité limite est solution de l'équation différentielle ordinaire

$$\frac{\partial}{\partial y}(b(y)q(y)) = \frac{1}{2} \frac{\partial^2}{\partial y^2}(\sigma^2(y)q(y)),$$

avec les conditions aux bords $q'(0) = q'(a) = 0$.

Nous pouvons alors résoudre cette équation et obtenir la mesure invariante du processus.

Chapitre 3

Dynamique des populations

Le hasard des événements viendra troubler sans cesse la marche lente, mais régulière de la nature, la retarder souvent, l'accélérer quelquefois. Condorcet (1743 - 1794), Esquisse d'un tableau historique des progrès de l'esprit humain.

Les processus de vie et de mort et les processus de branchement sont des modèles fondamentaux en dynamique des populations. Nous souhaitons ici étudier, à travers certains modèles probabilistes, les variations, la croissance, ou l'extinction d'une population. De nombreux domaines de biologie sont concernés par ces modèles : biologie des populations, cinétique des cellules, croissance bactériologique, réplication de l'ADN. Les questions posées sont les suivantes : pourquoi les extinctions de familles, de populations locales, d'espèces sont-elles si fréquentes ? Sont-elles conséquences de catastrophes ou de changements environnementaux externes aux populations, ou aussi liées à des comportements intrinsèques des populations ? Si tel est le cas, comment est-ce compatible avec la fameuse loi de croissance exponentielle malthusienne en cas de ressources importantes ? Y a-t-il une dichotomie entre croissance exponentielle et extinction ? ou des formes plus lentes de croissance ? Comment évolue la population si elle persiste en temps long ? Se stabilise-t-elle ?

Dans la suite, le terme de **population** désignera très généralement un ensemble d'individus, un **individu** pouvant tout aussi bien désigner une bactérie, un être humain porteur d'une certaine maladie, un éléphant...

Nous allons tout d'abord étudier les modèles à temps discret, décrivant principalement des dynamiques de généalogies, et rappeler en particulier les principaux résultats satisfaits par le processus de Bienaymé-Galton-Watson. Puis nous nous intéresserons aux processus de naissance et de mort en temps continu.

3.1 Processus de population en temps discret

3.1.1 Chaînes de Markov de vie et de mort

Dans les modèles en temps discret, le temps est en général exprimé en nombre de générations, mais peut aussi dans certains cas correspondre simplement à une discrétisation simple du temps continu. La variable X_n est l'effectif de la population à la génération n , et est donc à valeurs dans \mathbb{N} .

Nous étudierons le modèle général suivant. Si nous supposons qu'à un instant n , la population est de taille m ($X_n = m$), alors chaque individu vivant à la génération n

- meurt lorsque l'on passe de la génération n à la génération $n + 1$,
- donne alors naissance à un nombre aléatoire de descendants, **indépendamment** des autres individus.

On suppose de plus que les variables "nombre de descendants" ont une loi qui ne dépend pas de l'individu de la n -ième génération dont les descendants sont issus, mais qui peut éventuellement dépendre de m . On la note $Q^m = (q_i^m, i \in \mathbb{N})$. Ainsi, q_i^m est la probabilité qu'un individu issu d'une population de taille m donne naissance, à la génération suivante, à i individus.

- En outre, il y a possibilité d'immigration d'un nombre aléatoire d'individus, indépendamment du processus de naissance et mort, dont la loi $\eta^m = (\eta_j^m, j \in \mathbb{N})$ peut aussi dépendre de m . Ainsi η_j^m désigne la probabilité qu'un nombre j de migrants rejoigne une population de taille m à la génération suivante. Lorsque les migrants sont arrivés, ils se reproduisent et meurent comme les "anciens".

Construction de la chaîne : On se donne sur un espace (Ω, \mathcal{A}) une famille de variables aléatoires à valeurs entières

$$(X_0, (Y_{n,m,k})_{n,m,k \geq 0}, (Z_{n,m})_{n,m \geq 0}).$$

La variable aléatoire X_0 modélise le nombre d'individus au temps 0, $Y_{n,m,k}$ le nombre d'enfants du k -ième individus de la n -ième génération, si le nombre total d'individus de cette génération est m , et $Z_{n,m}$ modélise le nombre de migrants arrivant à la $(n + 1)$ -ième génération, sous la même hypothèse $X_n = m$.

Définition 3.1.1 *Pour chaque probabilité μ sur \mathbb{N} , on note \mathbb{P}_μ l'unique probabilité sur l'espace (Ω, \mathcal{A}) sous laquelle ces variables sont **indépendantes** et telles que*

- X_0 est de loi μ .
- Pour chaque n, m, k , la variable aléatoire $Y_{n,m,k}$ est de loi Q^m .
- Pour chaque n, m , la variable aléatoire $Z_{n,m}$ est de loi η^m .

Partant de X_0 , on définit donc les variables aléatoires X_n par récurrence :

$$X_{n+1} = \sum_{k=1}^m Y_{n,m,k} + Z_{n,m} \quad \text{si} \quad X_n = m, \quad \text{pour} \quad m \in \mathbb{N}, \quad (3.1.1)$$

avec la convention qu'une somme "vide" vaut 0.

On note \mathbb{P}_i la probabilité lorsque μ est la masse de Dirac en i : cela veut dire que l'on part de $X_0 = i$ individus.

On remarque que X_{n+1} est fonction de X_n et des variables $(Y_{n,m,k}, Z_{n,m}, m, k \geq 0)$. Il est alors clair, à cause des propriétés d'indépendance, que $(X_n)_n$ est une chaîne de Markov. La probabilité de transition est donnée par

$$P_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) = \sum_{k_1, k_2, \dots, k_i, r \in \mathbb{N}, k_1 + \dots + k_i + r = j} q_{k_1}^i q_{k_2}^i \cdots q_{k_i}^i \eta_r^i. \quad (3.1.2)$$

Définition 3.1.2 1) Lorsque les probabilités Q^i et η^i sont indépendantes de i , on dit que la dynamique de la population est **densité-indépendante**. Dans le cas contraire, on dira que la dynamique de la population est **densité-dépendante**.

2) On dit qu'il n'y a pas d'immigration lorsque $\eta_0^m = 1$, c'est à dire que η^m est la masse de Dirac en 0 pour tout m . La population est alors une **population isolée**.

On étudiera tout d'abord des modèles de population densité-indépendante et sans immigration. Dans ce cas, la loi de reproduction ne dépend pas du nombre d'individus présents. Ce modèle ne peut être réaliste que dans des situations de ressources infinies. En effet, dans ce cas, les individus ne sont pas en compétition pour survivre et chacun peut se reproduire librement et indépendamment des autres. En revanche, si les ressources (alimentaires par exemple) sont limitées, alors les individus vont entrer en compétition et le nombre d'individus de la population va avoir un effet sur la loi de reproduction. Cette densité-dépendance va introduire de la non-linéarité dans les modèles mathématiques.

On peut aussi compliquer le modèle de façon à décrire la dynamique de plusieurs sous-populations en interaction, ou en compétition. Donnons deux exemples.

Exemple 3.1.3 *Deux populations en interaction.* Dans cet exemple, on considère une population formée d'individus de deux types $a = 1$ ou $a = 2$. Il n'y a pas d'immigration. La génération n comprend $X_n^{(1)}$ individus de type 1, et $X_n^{(2)}$ individus de type 2. La loi de reproduction d'un individu de type 1, [resp 2], dans une population de m_1 individus de type 1 et m_2 individus de type 2, est la probabilité $Q_1^{m_1, m_2} = (Q_{1,i}^{m_1, m_2})_{i \in \mathbb{N}}$, [resp. $Q_2^{m_1, m_2} = (Q_{2,i}^{m_1, m_2})_{i \in \mathbb{N}}$].

Considérons alors une suite de variables aléatoires $(Y_{n, m_1, m_2, k}^a; n, m_1, m_2, k \in \mathbb{N}, a = 1 \text{ ou } 2)$ indépendantes entre elles et de loi $Q_a^{m_1, m_2}$. Introduisons également les effectifs initiaux $X_0^{(1)}$ de type 1 et $X_0^{(2)}$ de type 2.

On pose alors, pour $a = 1$ ou $a = 2$,

$$X_{n+1}^{(a)} = \sum_{k=1}^{m_a} Y_{n, m_1, m_2, k}^a \quad \text{si } X_n^{(1)} = m_1, X_n^{(2)} = m_2, \quad \text{pour tous } m_1, m_2 \in \mathbb{N}.$$

Le processus $(X_n^{(1)}, X_n^{(2)})$ est une chaîne de Markov à valeurs dans \mathbb{N}^2 , de matrice de transition

$$\begin{aligned} P_{(i_1, i_2)(j_1, j_2)} &= \mathbb{P}(X_{n+1}^{(1)} = j_1, X_{n+1}^{(2)} = j_2 | X_n^{(1)} = i_1, X_n^{(2)} = i_2) \\ &= \sum_{k_1 + \dots + k_{i_1} = j_1; r_1 + \dots + r_{i_2} = j_2} \prod_{l=1}^{i_1} Q_{1, k_l}^{i_1, i_2} \prod_{m=1}^{i_2} Q_{2, r_m}^{i_1, i_2}. \end{aligned}$$

Exemple 3.1.4 *Population isolée avec mutation.* La situation est la même que dans l'exemple 3.1.3, sauf que chaque descendant d'un individu de type 1 (resp. de type 2) peut "muter" et donc devenir de type 2 (resp. de type 1) avec la probabilité α_1 (resp. α_2), les mutations étant indépendantes de tout le reste. La chaîne $(X_n^{(1)}, X_n^{(2)})$ est encore markovienne, de matrice de transition donnée par

$$P_{(i_1, i_2)(j_1, j_2)} = \sum_{A(i_1, i_2, j_1, j_2)} \prod_{l=1}^{i_1} Q_{1, k_l}^{i_1, i_2} \binom{k_l}{k'_l} \alpha_1^{k'_l} (1 - \alpha_1)^{k_l - k'_l} \prod_{m=1}^{i_2} Q_{1, r_m}^{i_1, i_2} \binom{r_m}{r'_m} \alpha_2^{r'_m} (1 - \alpha_2)^{r_m - r'_m}$$

où $A(i_1, i_2, j_1, j_2)$ est l'ensemble des familles d'entiers $((k_l, k'_l, r_m, r'_m), l = 1, \dots, i_1; m = 1, \dots, i_2)$ telles que $0 \leq k'_l \leq k_l$ et $0 \leq r'_m \leq r_m$, avec $j_1 = (k_1 - k'_1) + \dots + (k_{i_1} - k'_{i_1}) + r'_1 + \dots + r'_{i_2}$ et $j_2 = k'_1 + \dots + k'_{i_1} + (r_1 - r'_1) + \dots + (r_{i_2} - r'_{i_2})$.

On peut encore compliquer les modèles, en supposant par exemple une reproduction sexuée, avec des individus mâles et femelles.

Les problèmes qui se posent alors naturellement concernent le comportement de la chaîne $(X_n)_n$ quand n tend vers l'infini. Par exemple, y a-t-il extinction de la population ($X_n \rightarrow 0$) ou un régime stationnaire (X_n converge en un certain sens) ou une "explosion" de la population ($X_n \rightarrow +\infty$), pour une population simple, ou extinction de l'un des types pour une population à deux types. Si $(X_n)_n$ tend vers l'infini, peut-on trouver une normalisation convenable (une suite a_n tendant vers 0) telle que $a_n X_n$ converge? Cela donne un ordre de grandeur de la taille de la population et de la vitesse de convergence. D'autres problèmes plus complexes peuvent se poser. Par exemple, que se passe-t-il lorsque l'on étudie le comportement de la population conditionnellement au fait qu'elle n'est pas éteinte à l'instant n ? C'est l'objet de la quasi-stationnarité qui donne un sens mathématique à des stabilités avant extinction de la population.

Nous n'allons pas pouvoir aborder ces problèmes dans toute leur généralité. Nous allons tout d'abord les développer dans le cadre le plus simple de chaîne de Bienaymé-Galton-Watson.

3.1.2 La chaîne de Bienaymé-Galton-Watson

La chaîne de vie et de mort la plus simple est la chaîne de Bienaymé-Galton-Watson que nous abrègerons en BGW. C'est une chaîne sans immigration, densité-indépendante. On

a

$$Q^m = Q \text{ pour tout } m$$

et η^m est la masse de Dirac en 0. Par suite, on étudie l'évolution du nombre X_n d'individus de cette population au cours des générations successives $n = 0, 1, 2, \dots$ en supposant que chacun des X_n individus de la n -ième génération engendre un nombre aléatoire $Y_{n,k}$ d'enfants ($1 \leq k \leq X_n$) de sorte que

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k} \quad (n \geq 0). \quad (3.1.3)$$

Les variables aléatoires $Y_{n,k}$ sont supposées indépendantes entre elles et de même loi. Cette loi, qui est la loi d'une variable générique Y , est une probabilité sur \mathbb{N} , définie par $(q_k, k \in \mathbb{N})$ et caractérisée par la fonction génératrice $g(s) = \mathbb{E}(s^Y) = \sum_k q_k s^k$. Dans toute la suite, nous excluons le cas $g(s) \equiv s$.

La figure 3.1 représente les individus des générations 0 à 3, lorsque $X_0 = 1$, $Y_{0,1} = 2$, $Y_{1,1} = 3$, $Y_{1,2} = 1$, $Y_{2,1} = 2$, $Y_{2,2} = 0$, $Y_{2,3} = 1$, $Y_{2,4} = 3$. Cette figure représente un arbre aléatoire.

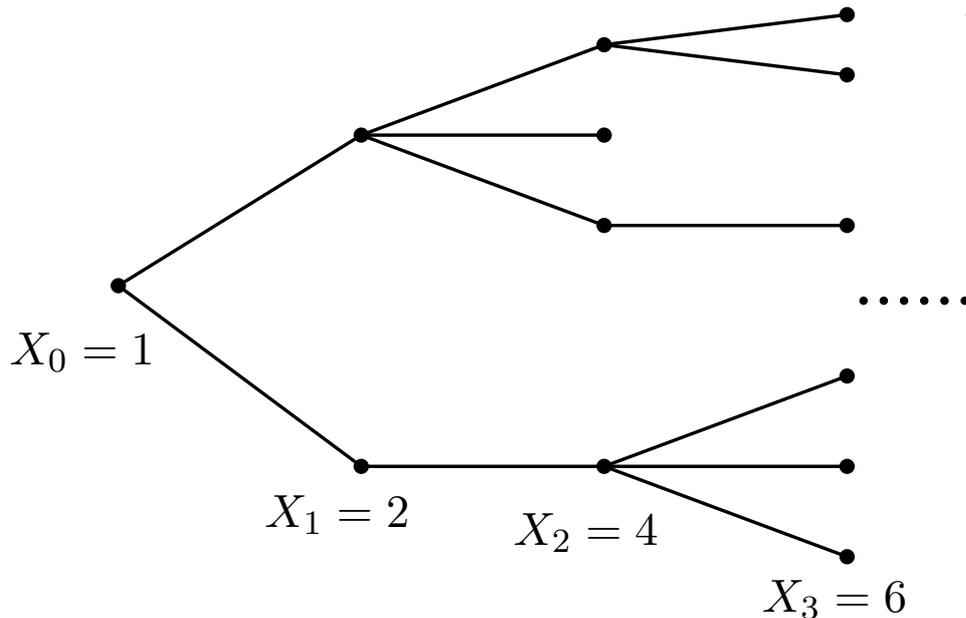


FIG. 3.1 – L'arbre généalogique du processus de branchement

Ce processus est le prototype le plus simple de processus de branchement, défini à temps discret et espace d'états discret. La première étude sur ces processus de branchement a été motivée par l'étude de la survivance des noms des familles des aristocrates anglais, et

plus précisément par l'étude de la non-extinction des descendants mâles qui seuls transmettaient le nom. Le modèle est ainsi le suivant : N hommes adultes d'une nation qui portent tous des noms de famille différents partent coloniser un pays. On suppose qu'à chaque génération, la proportion d'hommes qui ont k garçons est q_k , $k \in \mathbb{N}$.

De nombreux phénomènes de dynamique de populations peuvent en fait être modélisés par ce type de processus, dont la description est simple, mais qui présentent néanmoins un comportement non trivial. Le modèle décrit aussi bien la croissance d'une population de cellules, ou encore d'une épidémie dans une population, ou la réplication d'une molécule d'ADN en génétique. Considérons par exemple (voir Kimmel-Axelrod [15], Delmas-Jourdain) une macro-molécule d'ADN qui consiste en une chaîne de N nucléotides. En une unité de temps, cette chaîne est répliquée, chaque nucléotide étant copié de façon correcte avec probabilité p et indépendamment des autres nucléotides. A l'issue de la réplication, la molécule est détruite avec probabilité q ou bien donne naissance à deux molécules avec probabilité $1 - q$. On s'intéresse alors à la probabilité de disparition de la population de macro-molécules correctes. Elle est égale à la probabilité d'extinction d'un nom de famille dans le cas où $q_0 = q$ (destruction), $q_1 = (1 - q)(1 - p^N)$ (non destruction mais réplication incorrecte), $q_2 = (1 - q)p^N$ (non destruction et réplication correcte), et $q_k = 0$ pour $k \geq 3$.

Résultats élémentaires

Une propriété immédiate, mais fondamentale, est que 0 est un état absorbant : si $X_n = 0$, alors $X_{n+1} = 0$. De manière équivalente, le temps d'atteinte de 0 , c'est à dire

$$T = \inf\{n \geq 0, X_n = 0\}$$

est appelé **temps d'extinction**, au sens où $X_n = 0$ si et seulement si $n \geq T$.

Un problème biologique important sera de connaître la loi du temps d'extinction, et en particulier la probabilité d'extinction partant d'un nombre i d'individus :

$$p_i = \mathbb{P}_i(T < +\infty)$$

pour $i \geq 1$. (bien-sûr, $p_0 = 1$).

Une autre propriété importante de cette chaîne est la *propriété de branchement*. Une chaîne de Markov $(X_n)_n$ à valeurs entières a cette propriété si sa loi (en tant que processus) sous \mathbb{P}_i est la même que la loi de

$$(Z_n^1 + \dots + Z_n^i)_n$$

où les $(Z_n^j)_n, j \geq 1$ sont des chaînes de markov indépendantes entre elles, et qui ont toutes même loi que la chaîne $(X_n)_n$ sous \mathbb{P}_1 . En vertu de (3.1.3), c'est bien le cas ici. En effet, si $X_0 = i$, on peut en fait écrire $X_n = \sum_{j=1}^i Z_n^j$, où Z_n^j représente le nombre d'individus vivant à l'instant n , et descendant du j -ième individu initial.

Remarquons d'ailleurs que la propriété de branchement pour une chaîne de Markov est caractéristique des chaînes BGW. En effet, la propriété de branchement implique en particulier que la loi $(P_{ij}, j \in \mathbb{N})$ est la puissance i -ième de convolution de la loi $(P_{1j}, j \in \mathbb{N})$. En d'autres termes, on a (3.1.2) avec $q_j^m = p_{1j}$ et $\eta_0^m = 1$. La chaîne est donc BGW. Plus généralement, la propriété de branchement permet de ramener l'étude de \mathbb{P}_i à celle de \mathbb{P}_1 . Par exemple,

- La loi $(P_{ij}^n)_{j \in \mathbb{N}}$ est la puissance i -ième de convolution de la loi $(P_{1j}^n)_{j \in \mathbb{N}}$
- On a

$$\mathbb{P}_i(T \leq n) = \mathbb{P}_1(T \leq n)^i. \quad (3.1.4)$$

Pour ce dernier fait, on remarque que le temps d'extinction de $Z_n^1 + \dots + Z_n^i$ est le maximum des temps d'extinction des $Z_n^j, 1 \leq j \leq i$.

D'un point de vue des propriétés de la chaîne, et en particulier du temps d'extinction, un certain nombre de cas sont triviaux :

- Si $q_1 = 1$, on a $X_n = X_0$ et l'effectif de la population ne varie pas.
- Si $q_1 < 1$ et si $q_0 + q_1 = 1$, on a $X_{n+1} \leq X_n$. L'effectif de la population décroît. On a $T = 1$ \mathbb{P}_1 -p.s. si $q_0 = 1$ et $\mathbb{P}_1(T = k) = q_0 q_1^{k-1}$ si $q_0 < 1$. (T suit une loi géométrique).
- Si $q_1 < 1$ et $q_0 = 0$, on a $X_{n+1} \geq X_n$, l'effectif de la population croît, et on a $T = +\infty$ \mathbb{P}_1 -p.s.

Dans la suite on élimine ces 3 cas en supposant que

$$q_0 > 0, \quad q_0 + q_1 < 1. \quad (3.1.5)$$

Proposition 3.1.5 *L'état 0 est absorbant, et sous (3.1.5), les autres états sont transients. Si de plus d est le PGCD des $i \geq 1$ tels que $q_i > 0$, les classes sont $C = d\mathbb{N}^* = \{nd, n \geq 1\}$, plus tous les singletons $\{i\}$ avec $i \notin C$.*

Preuve. D'après (3.1.3), si $i \geq 1$, on a $p_{i0} \geq (q_0)^i > 0$, donc i mène à 0, alors que 0 ne mène pas à $i \geq 1$ puisque 0 est absorbant. Donc l'état i ne peut pas être récurrent. Les $Y_{n,k}$ prennent p.s. leurs valeurs dans $\{i, q_i > 0\}$, qui est contenu dans $C \cup \{0\}$, de sorte que pour tout $n \geq 1$, $X_n \in C \cup \{0\}$ p.s.. Par suite, un état i ne peut mener qu'à 0 et aux points de C et tout singleton non contenu dans C est une classe. \square

Du point de vue des propriétés ergodiques ou de la stationnarité éventuelle de la chaîne $(X_n)_n$, tout est dit dans la proposition précédente : il y a une seule classe de récurrence $\{0\}$ et une seule probabilité invariante, la masse de Dirac en 0. La chaîne est stationnaire si et seulement si elle part de 0 et elle reste alors toujours en 0. Sinon, du fait que tous les états non nuls sont transients, la chaîne n'a que deux comportements possibles, soit elle converge vers 0, soit elle converge vers $+\infty$.

Du point de vue des applications en revanche, on va répondre à un certain nombre de questions : le calcul des probabilités d'extinction p_i , la loi du temps d'extinction T , la vitesse d'explosion si $(X_n)_n$ tend vers l'infini, etc.

Le comportement à l'infini

Pour les chaînes BGW, nous allons pouvoir décrire le comportement presque-sûr de la chaîne, quand n tend vers l'infini. L'outil principal sera la fonction génératrice g de la loi Q définie ci-dessous sur $[0, 1]$ par

$$g(s) = \sum_{k=0}^{\infty} q_k s^k = \mathbb{E}(s^Y). \quad (3.1.6)$$

On introduit également les deux premiers moments de cette loi (qui peuvent éventuellement être infinis)

$$m = \sum_{k=0}^{\infty} k q_k ; m_2 = \sum_{k=0}^{\infty} k^2 q_k. \quad (3.1.7)$$

On sait que g est croissante convexe. Elle est de classe C^∞ sur $[0, 1[$, et est une fois (resp. deux fois) dérivable à gauche en $s = 1$ si et seulement si $m < \infty$ (resp. $m_2 < \infty$), et dans ce cas, $m = g'(1)$ (resp. $m_2 - m = g''(1)$). (Pour les propriétés de la fonction génératrice g , vous pouvez vous reporter au polycopié de 1ère année).

On note g_n la n -ième itérée de g , c'est à dire $g \circ g \circ \dots \circ g$, n fois. On a $g_{n+1} = g \circ g_n = g_n \circ g$. Enfin, on pose

$$s_0 = \inf\{s \in [0, 1], g(s) = s\}. \quad (3.1.8)$$

On a $s_0 > 0$ car $q_0 > 0$ et $s_0 \leq 1$ car $g(1) = 1$.

On a alors le

- Lemme 3.1.6** a) Si $g'(1) \leq 1$, alors $s_0 = 1$ et $g_n(s)$ croît vers 1 lorsque $n \nearrow \infty$ pour tout $s \in [0, 1]$.
 b) Si $g'(1) > 1$, l'équation $g(v) = v$ possède une solution unique s_0 dans $[0, 1[$ et $g_n(s)$ croît vers s_0 , [resp. décroît], lorsque $n \nearrow \infty$, pour tout $s \in [0, v]$, [resp. tout $s \in [v, 1[$].

Preuve. L'application $s \rightarrow g(s)$ de l'intervalle $[0, 1]$ dans lui-même est croissante et strictement convexe. De plus $g(1) = 1$. Comme nous avons exclu le cas $g(s) \equiv s$, la courbe g ne coupe pas ou au contraire coupe la diagonale du carré $[0, 1]^2$ en un point distinct de $(1, 1)$, selon que $g'(1) \leq 1$ ou au contraire que $g'(1) > 1$; ceci se voit bien sur la figure 3.2. Ainsi, selon le cas, l'équation de point fixe $g(v) = v$ n'a pas de solution ou au contraire possède une unique solution dans $[0, 1[$, c'est à dire que soit $s_0 = 1$, soit $s_0 < 1$.

a) Lorsque $g'(1) \leq 1$, nous avons $s \leq g(s)$ et donc $g_n(s) \leq g_{n+1}(s)$ (puisque $g_{n+1} = g \circ g_n$) pour tout s . La limite $\lim_n g_n(s)$ est inférieure à 1 et solution de $g(u) = u$, elle ne peut alors valoir que 1.

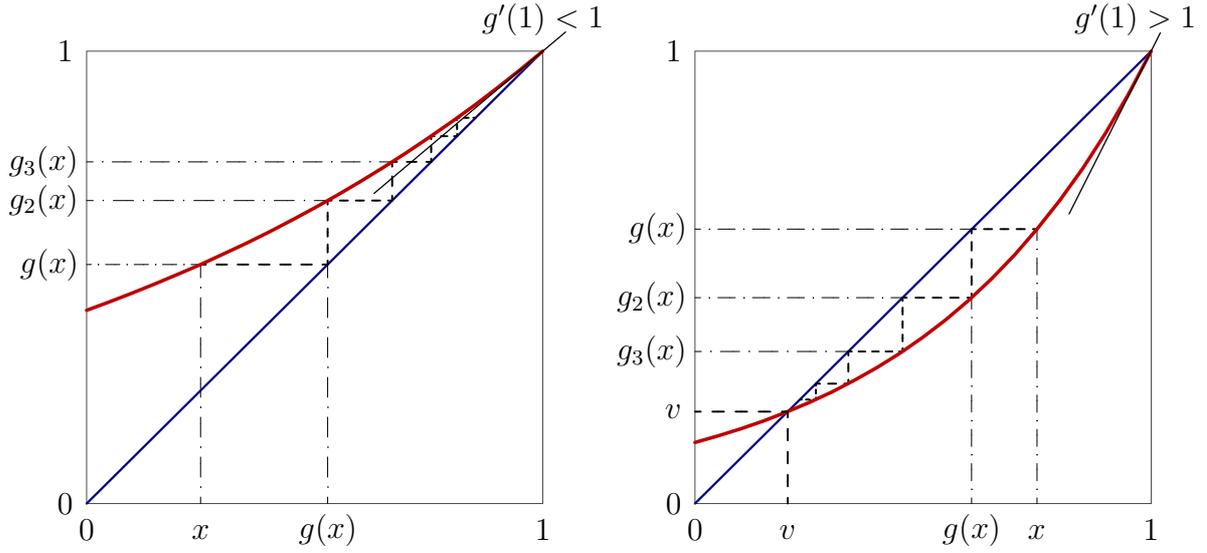


FIG. 3.2 – La fonction g et ses itérées dans les deux cas $m \leq 1$ et $m > 1$ ($v = s_0$).

b) De même, si $g'(1) > 1$, nous avons $s \leq g(s) \leq s_0$ ou $s \geq g(s) \geq s_0$ selon que $s \leq s_0$ ou que $s \geq s_0$; il s'en suit que $g_n(s)$ croît (resp. décroît) avec n selon le cas. La limite $\lim_n g_n(s)$, qui est solution de $g(u) = u$ et est strictement inférieure à 1, (du moins si $s \neq 1$), est alors nécessairement égale à s_0 . \square

On a aussi les propriétés suivantes.

Lemme 3.1.7 1) On a $1 - g_n(s) \leq m^n(1 - s)$.

2) Si $m = 1$ et $m_2 < \infty$, alors pour tout $s \in [0, 1[$, la suite $1 - g_n(s)$ est équivalente à $\frac{2}{n(m_2-1)}$ quand $n \rightarrow \infty$, pour tous $s \in [0, 1[$.

Preuve. L'inégalité $1 - g(s) \leq m(1 - s)$ est évidente, puisque g' est croissante sur $[0, 1]$ et que $g'(1) = m$. En itérant cette inégalité, on obtient immédiatement l'assertion 1).

Posons $a = \frac{m_2-1}{2}$. On a $a > 0$ (car $m = 1$), et un développement au voisinage de $s = 1$ montre que pour tout $s \in [0, 1[$,

$$1 - g(s) = (1 - s)(1 - (1 - s)(a + \gamma(s))), \quad \text{où } \lim_{s \rightarrow 1} \gamma(s) = 0,$$

puisque $g'(1) = 1$ et $g''(1) = m_2 - m = 2a$. Par suite,

$$\frac{1}{1 - g(s)} = \frac{1}{1 - s} + a + \Gamma(s), \quad \text{où } \lim_{s \rightarrow 1} \Gamma(s) = 0.$$

Nous en déduisons que

$$\frac{1}{1 - g_{j+1}(s)} = \frac{1}{1 - g_j(s)} + a + \Gamma(g_j(s)),$$

et en réitérant ces égalités pour $j = n - 1, \dots, 1$, on obtient finalement que

$$\frac{1}{1 - g_n(s)} = \frac{1}{1 - s} + na + \Delta_n(s), \quad \text{où } \Delta_n(s) = \sum_{j=0}^{n-1} \Gamma(g_j(s)).$$

Pour $s \in [0, 1[$ fixé, on sait que $g_j(s) \rightarrow 1$ donc $\Gamma(g_j(s)) \rightarrow 0$ quand $j \rightarrow \infty$, puisque $m = 1$. On en déduit que $\frac{\Delta_n(s)}{n} \rightarrow 0$ quand n tend vers l'infini, comme limite de Césaro de la suite $(\Gamma(g_n(s)))_n$. Ainsi, quand n tend vers l'infini, on a

$$1 - g_n(s) = \frac{1}{na} \frac{1}{1 + \frac{1}{na(1-s)} + \frac{\Delta_n(s)}{na}} \sim \frac{1}{na} = \frac{2}{n(m_2 - 1)}.$$

On a donc montré 2). □

Proposition 3.1.8 *Soit G_n la fonction génératrice de X_n . Alors $\forall n \geq 0$,*

$$G_{n+1}(s) = G_n(g(s)). \quad (3.1.9)$$

En particulier, si nous supposons que $X_0 = i$, on a

$$G_n(s) = g_n(s)^i \quad (3.1.10)$$

$$\mathbb{E}_i(X_n) = \mathbb{E}(X_n | X_0 = i) = m^n i. \quad (3.1.11)$$

Preuve. On a

$$\begin{aligned} G_{n+1}(s) &= \sum_k \mathbb{E}(s^{\sum_{i=1}^k Y_{n,i}} \mathbf{1}_{\{X_n=k\}}) \\ &= \sum_k (g(s))^k \mathbb{P}(X_n = k) = G_n(g(s)), \end{aligned}$$

par indépendance des $Y_{n,k}$ entre eux et avec X_n .

(3.1.10) en découle facilement, puis en dérivant cette expression et en faisant tendre s vers 1, on obtient (3.1.11). □

Ce résultat va nous permettre d'étudier le comportement asymptotique de la suite $(X_n)_{n \geq 0}$ et de calculer notamment la probabilité d'extinction de la population.

Remarquons que $X_{n+1} = 0$ si $X_n = 0$ et donc les événements $\{X_n = 0\}$ croissent avec n . L'événement $A =$ "Extinction de la population" est donc naturellement défini par

$$A = \{T < +\infty\} = \cup_n \uparrow \{X_n = 0\} \quad (3.1.12)$$

et sa probabilité donnée par

$$\mathbb{P}(A) = \lim_n \uparrow G_n(0).$$

L'étude la suite $(G_n(0))_n$ va nous amener à distinguer 3 cas :

- Le cas **sous-critique**, quand $m < 1$,
- Le cas **critique**, quand $m = 1$,
- Le cas **surcritique**, quand $m > 1$.

Théorème 3.1.9 1) Dans les cas sous-critique et critique, ($m \leq 1$), on a $p_i = \mathbb{P}_i(T < \infty) = 1$. La population s'éteint presque sûrement.

2) Dans le cas sous-critique, on a $\mathbb{E}_i(T) < \infty$, tandis que dans le cas critique, on a $\mathbb{E}_i(T) = \infty$.

3) Dans le cas surcritique ($m > 1$), sous \mathbb{P}_i , la suite X_n converge presque-sûrement vers une limite X_∞ qui ne prend que les valeurs 0 et $+\infty$ avec les probabilités

$$p_i = \mathbb{P}_i(T < \infty) = \mathbb{P}_i(X_\infty = 0) = s_0^i, \quad \mathbb{P}_i(X_\infty = +\infty) = 1 - s_0^i,$$

où $s_0 \in]0, 1[$ a été défini en (3.1.8).

Remarque 3.1.10 Par indépendance des variables $(Y_{n,k}; n \geq 1, k \geq 1)$ on déduit de (3.1.3) que $\mathbb{E}(X_n) = \mathbb{E}(Y) \times \mathbb{E}(X_{n-1}) = (\mathbb{E}(Y))^n$ (par récurrence). Cette dernière égalité entraîne que la population s'éteint lorsque $m = \mathbb{E}(Y) < 1$. En effet, il est facile de voir que

$$\mathbb{P}_i(X_n \neq 0) \leq \mathbb{E}_i(X_n) = m^n i. \quad (3.1.13)$$

On va donner ci-dessous une preuve qui regroupe les 3 cas.

Preuve. La démonstration du théorème repose sur le lemme 3.1.6.

1) On a $\{T \leq n\} = \{X_n = 0\}$, et donc

$$\mathbb{P}_i(T \leq n) = \mathbb{P}_1(T \leq n)^i = g_n(0)^i \quad (3.1.14)$$

qui converge vers s_0^i quand n tend vers l'infini. Ainsi donc, cela entraîne que $\mathbb{P}_i(T < \infty) = 1$ si $m \leq 1$ d'où la première partie du théorème. De plus, $\mathbb{P}_i(T < \infty) = s_0^i$ quand $m > 1$.

2) On a aussi

$$\mathbb{E}_i(T) = \sum_{n \geq 0} \mathbb{P}_i(T > n) = \sum_{n \geq 0} (1 - g_n(0)^i).$$

Le lemme 3.1.7 entraîne immédiatement que si $m < 1$, alors $\mathbb{E}_1(T) < \infty$, puisque dans ce cas, la série de terme général m^n est convergente. De plus, nous avons vu que le temps d'extinction de X_n sachant que $X_0 = i$ est le maximum des temps d'extinction des sous-populations issues de chaque individu. Ainsi, pour tout $i \in \mathbb{N}^*$, $\mathbb{E}_i(T) < \infty$.

Si $m = 1$ et $m_2 < \infty$, alors par la deuxième partie du lemme, on sait que $\mathbb{E}_1(T) = \infty$ (la série harmonique diverge), ce qui entraîne alors que $\mathbb{E}_i(T) = \infty$ pour tout $i \geq 1$. On a donc montré l'assertion (2) du théorème. Si $m_2 = \infty$, la preuve est plus délicate. Toutefois

on peut accepter l'idée intuitive suivante, qui peut se justifier mathématiquement. La loi de branchement va charger plus les grandes valeurs entières et l'espérance du temps d'extinction devrait être plus grande que dans le cas où m_2 est fini, et donc $\mathbb{E}_i(T) = \infty$.

3) Il reste à étudier le comportement asymptotique de X_n quand $m > 1$.

Lorsque $m > 1$, l'extinction n'est pas certaine puisque $\mathbb{P}(A) = s_0 < 1$. Rappelons que l'état 0 est absorbant et que les autres états sont transients. Donc deux cas seulement sont possibles : ou bien $X_n = 0$ à partir d'un certain rang (c'est l'extinction), ou bien $X_n \rightarrow \infty$ (c'est l'explosion). Ces deux événements complémentaires ont pour probabilités respectives s_0^i et $1 - s_0^i$. Cela conclut la preuve du théorème. \square

La formule (3.1.14) donne en fait la fonction de répartition du temps d'extinction T , et donc sa loi, sous \mathbb{P}_1 et sous \mathbb{P}_i . On a :

$$\mathbb{P}_1(T = n) = \begin{cases} 0 & \text{si } n = 0 \\ g_n(0) - g_{n-1}(0) & \text{si } n \in \mathbb{N}^* \\ 1 - s_0 & \text{si } n = +\infty. \end{cases} \quad (3.1.15)$$

Remarquons que dans le cas surcritique, T est une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

Remarque 3.1.11 Nous avons montré que, dans le cas où la population initiale est composée de i individus, nous nous trouvons dans l'un des 3 cas suivants.

- Lorsque $m < 1$, la probabilité de " non-extinction à l'instant n " tend vers 0 à une vitesse géométrique puisque, comme on l'a vu,

$$\mathbb{P}_i(X_n \neq 0) \leq m^n i.$$

- Lorsque $m = 1$, cette même probabilité tend beaucoup plus lentement vers zéro. Plus précisément, a vu que si $m_2 < \infty$,

$$\mathbb{P}_i(X_n \neq 0) \sim \frac{2i}{n(m_2 - 1)} \quad \text{lorsque } n \rightarrow \infty.$$

- Lorsque $m > 1$, X_n tend vers 0 ou vers $+\infty$.

Dans le cas surcritique, il est intéressant de préciser à quelle vitesse X_n tend vers l'infini, sur l'ensemble $\{T = \infty\}$, qui est donc de probabilité strictement positive $1 - s_0$. Nous nous limitons au cas $m_2 < \infty$.

Théorème 3.1.12 Si $m > 1$ et $m_2 < \infty$, la suite $\frac{X_n}{m^n}$ converge \mathbb{P}_i -p.s. vers une variable aléatoire U qui est presque-sûrement strictement positive sur l'ensemble $\{T = \infty\}$, et telle que $E_i(U) = i$.

Ce résultat nous dit que sur l'ensemble où il n'y a pas extinction, la croissance de X_n vers $+\infty$ est exponentielle (en m^n). C'est l'expression probabiliste de **la loi de Malthus**.

Preuve. Posons $U_n = \frac{X_n}{m^n}$. Considérons la filtration $(\mathcal{F}_n)_n$ formée de la suite croissante de tribus \mathcal{F}_n , où \mathcal{F}_n est la tribu engendrée par les $(X_k, k \leq n)$. Comme m et m_2 sont les deux premiers moments des variables $Y_{n,k}$, on a en utilisant (3.1.3) que

$$E_i(X_{n+1}|\mathcal{F}_n) = mX_n ; E_i(X_{n+1}^2|\mathcal{F}_n) = m_2X_n + m^2X_n(X_n - 1), \quad (3.1.16)$$

et donc

$$E_i(U_{n+1}|\mathcal{F}_n) = U_n ; E_i(U_{n+1}^2|\mathcal{F}_n) \leq U_n^2 + \frac{m_2}{m^{n+2}}U_n. \quad (3.1.17)$$

Par suite, $E_i(U_n) = i$ et, comme $m > 1$,

$$E_i(U_{n+1}^2) \leq E_i(U_n^2) + \frac{im_2}{m^{n+2}} \leq i^2 + \frac{im_2}{m^2} \sum_{j=0}^n m^{-j} \leq i^2 + \frac{im_2}{m(m-1)}.$$

On en déduit finalement que U_n est une martingale, dont les moments d'ordre 2 sont uniformément bornés. Elle converge donc \mathbb{P}_i -presque-sûrement et en moyenne vers U . (Voir le cours de MAP 432). En particulier, $\mathbb{E}_i(U) = i$. Il reste à montrer que $U > 0$, \mathbb{P}_i -p.s. sur $\{T = \infty\}$. On pose $r_i = \mathbb{P}_i(U = 0)$. En vertu de la propriété de branchement, la loi de U sous \mathbb{P}_i est la même que celle de $U^1 + \dots + U^i$, où les U^j sont des variables aléatoires indépendantes de même loi que U sous \mathbb{P}_1 . Par suite, on a $r_i = (r_1)^i$. Par la propriété de Markov, on a

$$\mathbb{P}_1(U = 0|\mathcal{F}_1) = \mathbb{P}_{X_1}(U = 0) = r_1^{X_1}.$$

En prenant l'espérance, on obtient $r_1 = \mathbb{E}_1(r_1^{X_1}) = g(r_1)$. Ainsi, r_1 est solution de l'équation $g(s) = s$, équation qui dans $[0, 1]$ admet les deux solutions s_0 et 1. Or $\mathbb{E}_1(U) = 1$ implique que $r_1 = \mathbb{P}_1(U = 0) < 1$, donc nécessairement $r_1 = s_0$ et $r_i = s_0^i$, c'est à dire que $\mathbb{P}_i(U = 0) = \mathbb{P}_i(T < \infty)$. Mais à l'évidence, $\{T < \infty\} \subset \{U = 0\}$, donc nécessairement $\{U > 0\}$ \mathbb{P}_i -p.s. sur $\{T = \infty\}$. \square

Relation entre processus de BGW et modèle de généalogie

Etudions le comportement d'un processus de BGW conditionné à rester de taille constante N . Nous allons voir que dans ce cas, pour une loi de reproduction du BGW bien précise, la loi conditionnelle est liée au modèle de Wright-Fisher, que nous étudierons en détail dans le chapitre "Généétique des populations".

Théorème 3.1.13 *Considérons un processus X de BGW, dont la loi de reproduction est une loi de poisson de paramètre m . Alors, conditionnellement au fait que X garde la même valeur N , la répartition des descendants de chaque individu de la génération précédente (Y_1, \dots, Y_N) suit une loi multinomiale de paramètres $(N, \frac{1}{N}, \dots, \frac{1}{N})$.*

Preuve. Soit D_i le nombre de descendants de l'individu i dans le modèle de BGW conditionné à rester de taille N . Soit (k_1, \dots, k_N) , un N -uplet d'entiers de somme N . On a alors

$$\begin{aligned} \mathbb{P}(D_1 = k_1, \dots, D_N = k_N) &= \mathbb{P}(Y_1 = k_1, \dots, Y_N = k_N | X_1 = N) \\ &= \frac{\prod_{i=1}^N \mathbb{P}(Y_i = k_i)}{\mathbb{P}_N(X_1 = N)}, \end{aligned}$$

car les variables aléatoires Y_i sont indépendantes. Appelons comme précédemment g la fonction génératrice de la loi de reproduction, qui est supposée être une loi de Poisson de paramètre m . On a pour tout $s \in [0, 1]$

$$g(s) = \exp(-m(1-s)),$$

et la fonction génératrice de X_1 sous \mathbb{P}_N est $g(s)^N = \exp(-Nm(1-s))$. La loi de X_1 sous \mathbb{P}_N est donc une loi de Poisson de paramètre mN . Nous en déduisons que

$$\begin{aligned} \mathbb{P}(D_1 = k_1, \dots, D_N = k_N) &= \prod_{i=1}^N e^{-m} \frac{m^{k_i}}{k_i!} \frac{N!}{e^{-mN} (Nm)^N} \\ &= \frac{N!}{k_1! \cdots k_N!} \left(\frac{1}{N}\right)^N. \end{aligned}$$

Nous reconnaissons ici la loi multinomiale de paramètres $(N, \frac{1}{N}, \dots, \frac{1}{N})$. □

3.1.3 Chaîne BGW avec immigration

On vient de voir que la chaîne BGW a un comportement limite "dégénéré" en un certain sens, et il n'y a pas de probabilité invariante hormis la masse de Dirac en 0. La situation devient différente, avec notamment l'existence possible de probabilités invariantes non triviales, lorsqu'on ajoute une immigration.

Nous reprenons donc le modèle introduit au Chapitre 3.1.1, en supposant que pour tout m ,

$$Q^m = Q \quad ; \quad \eta^m = \eta.$$

Ainsi, la chaîne de vie et de mort densité-indépendante avec immigration est décrite par

$$X_{n+1} = \sum_{k=1}^m Y_{n,k} + Z_n \quad \text{si } X_n = m, \quad \text{pour } m \in \mathbb{N}. \quad (3.1.18)$$

Les variables aléatoires $(Z_n)_n$ ont la même loi η qu'une variable aléatoire Z .

La différence essentielle avec la chaîne simple (sans immigration) est que 0 n'est plus un état absorbant, puisque $p_{0i} = \eta_i$ est strictement positif pour au moins un $i \geq 1$.

Certains cas sont inintéressants :

- Si $\eta_0 = 1$, on se ramène au cas précédent.
- Si $q_0 = 1$, $X_{n+1} = Z_n$ et la chaîne se réduit à une suite de variables aléatoires indépendantes et de même loi.
- Si $q_0 = 0$, on a $X_{n+1} \geq X_n + Z_n$ et la population ne peut avoir un temps d'extinction fini.

Dans la suite, on suppose donc que

$$0 < q_0 < 1, \quad q_1 < 1, \quad \eta_0 < 1. \quad (3.1.19)$$

On gardera la notation m du paragraphe précédent, qui désigne l'espérance de la loi de reproduction. Il faut bien comprendre que chaque immigrant, une fois arrivé, se comporte de la même manière que les autres individus et donc engendre un arbre de BGW composé de ses descendants, avec la même loi de reproduction.

Les théorèmes suivants décrivent alors le comportement du processus de BGW avec immigration. Ces théorèmes généralisent le cas sans immigration, et sont plus subtils à montrer. Nous ne donnerons que des démonstrations partielles.

Théorème 3.1.14 *Supposons $m < 1$. Notons $\ln^+ Z = \sup(\ln Z, 0)$. On a alors la dichotomie suivante.*

$$\begin{aligned} \mathbb{E}(\ln^+ Z) < \infty &\Rightarrow (X_n) \text{ converge en loi.} \\ \mathbb{E}(\ln^+ Z) = \infty &\Rightarrow (X_n) \text{ converge en probabilité vers } +\infty. \end{aligned}$$

Remarquons que ce théorème entraîne en particulier la convergence de $(X_n)_n$ si $m < 1$ et $\mathbb{E}(Z) < \infty$.

Théorème 3.1.15 *Supposons $m > 1$. On a alors la dichotomie suivante.*

$$\begin{aligned} \mathbb{E}(\ln^+ Z) < \infty &\Rightarrow \lim_n m^{-n} X_n \text{ existe et est finie p.s.} \\ \mathbb{E}(\ln^+ Z) = \infty &\Rightarrow \limsup_n c^{-n} X_n = \infty \text{ pour toute constante } c > 0. \end{aligned}$$

Pour prouver ces théorèmes, nous allons utiliser le lemme suivant :

Lemme 3.1.16 *Soient ζ_1, ζ_2, \dots des variables aléatoires indépendantes et de même loi distribuée comme ζ . Alors pour tout $c > 1$,*

$$\begin{aligned} \mathbb{E}(\ln^+ \zeta) < \infty &\Rightarrow \sum_{n \geq 1} c^{-n} \zeta_n < \infty \text{ p.s.} \\ \mathbb{E}(\ln^+ \zeta) = \infty &\Rightarrow \limsup_n c^{-n} \zeta_n = \infty \text{ p.s.} \end{aligned}$$

Preuve du lemme Par le théorème de Borel-Cantelli, on peut prouver que pour toute suite de v.a. W_1, W_2, \dots positives indépendantes et équidistribuées (de même loi que W),

$$\limsup_n \frac{W_n}{n} = 0 \text{ ou } \infty$$

suivant que $\mathbb{E}(W)$ est fini ou non. En effet, soit $\varepsilon > 0$. On a alors

$$\sum_n \mathbb{P}\left(\frac{W_n}{n} > \varepsilon\right) = \sum_n \sum_{i > n\varepsilon} \mathbb{P}(W_n = i) = \sum_i \sum_{n < \frac{i}{\varepsilon}} \mathbb{P}(W_n = i) = \sum_i \frac{i}{\varepsilon} \mathbb{P}(W = i) = \frac{1}{\varepsilon} \mathbb{E}(W).$$

Ainsi, si $\mathbb{E}(W) < \infty$, par le Théorème de Borel-Cantelli (voir Polycopié MAP 311), on a que $\mathbb{P}(\limsup_n \{\frac{W_n}{n} > \varepsilon\}) = 0$, ce qui entraîne que la suite $(\frac{W_n}{n})_n$ converge p.s. vers 0. Si au contraire $\mathbb{E}(W) = \infty$, et puisque les W_i sont indépendantes, alors la probabilité ci-dessus vaut 1, et donc $\frac{W_n}{n}$ tend vers l'infini.

En utilisant ce résultat, nous montrons le lemme. Remarquons que si les résultats asymptotiques annoncés sont vrais pour la sous-suite des $\zeta_n > 0$, ils seront encore vrais pour tous les ζ_n . On pose alors $W_n = \ln^+(\zeta_n)$ et on définit $a := \ln c > 0$. Alors on a $c^{-n}\zeta_n = \exp -n(a - \frac{W_n}{n})$ tant que $\zeta_n > 1$.

Preuve. (du Théorème 3.1.14). Remarquons que le processus de BGW avec immigration X issu de 0 et évalué à la génération n vérifie

$$X_n \text{ a même loi que } \sum_{k=0}^n \xi_k,$$

où les ξ_k sont indépendants et pour chaque k , ξ_k décrit la contribution à la génération n des immigrants arrivés à la génération $n - k$. Ainsi ξ_k décrit un processus de BGW sans immigration lié à la loi de reproduction Y , à la génération k .

Nous noterons ζ_k la condition initiale de ξ_k , c'est-à-dire le nombre d'immigrants à la génération $n - k$. Remarquons que $\xi_0 = \zeta_n$, ξ_1 est égal au nombre d'enfants des ζ_{n-1} migrants de la génération $n - 1$, etc. De plus, chaque ζ_k a même loi que Z , donc

$$\mathbb{E}(\ln^+ Z) = \mathbb{E}(\ln^+ \zeta).$$

Nous cherchons donc à déterminer si

$$X_\infty := \sum_{k=1}^{\infty} \xi_k,$$

avec ξ_k tous indépendants, est fini ou infini. On peut montrer (grâce à ce qu'on appelle la loi du tout ou rien), que X_∞ est fini p.s. ou infini p.s. Soit \mathcal{G} la tribu engendrée par toutes les variables $\zeta_k, k \in \mathbb{N}$. Supposons tout d'abord que $\mathbb{E}(\ln^+ Z) < \infty$. On a alors

$$\mathbb{E}(X_\infty | \mathcal{G}) = \sum_{k=0}^{\infty} \zeta_k m^k.$$

En effet ξ_0 a pour loi ζ , l'espérance conditionnelle de ξ_1 sachant \mathcal{G} est égale à l'espérance du nombre d'individu de la première génération issue de ζ_1 individus, c'est-à-dire $\zeta_1 m$, etc. On peut alors appliquer le lemme 3.1.16 (puisque $m < 1$), et on en déduit que l'espérance conditionnelle est finie p.s., ce qui entraîne que X_∞ est fini p.s.

Par un argument un peu plus compliqué que nous ne développerons pas ici, on peut montrer la réciproque, à savoir que si X_∞ est fini p.s., alors $\mathbb{E}(\ln^+(Z)) < \infty$. \square

Preuve.(du Théorème 3.1.15). Considérons tout d'abord le cas où $\mathbb{E}(\ln^+ Z) = \infty$. Alors grâce au lemme 3.1.16, on sait que $\limsup_n c^{-n} Z_n = \infty$, pour tout $c > 1$. Puisque $X_n \geq Z_n$, on en déduit le résultat.

Supposons maintenant que $\mathbb{E}(\ln^+ Z) < \infty$. On appelle \mathcal{F}_n la tribu engendrée par X_0, X_1, \dots, X_n ainsi que par toutes les variables aléatoires $Z_k, k \in \mathbb{N}$. On a alors

$$\begin{aligned} \mathbb{E}\left(\frac{X_{n+1}}{m^{n+1}} \mid \mathcal{F}_n\right) &= \frac{1}{m^{n+1}} \mathbb{E}\left(\sum_{i=1}^{X_n} Y_i + Z_{n+1} \mid \mathcal{F}_n\right) \\ &= \frac{X_n}{m^n} + \frac{Z_{n+1}}{m^{n+1}}. \end{aligned}$$

Nous en déduisons que $(\frac{X_n}{m^n})_n$ est une sous-martingale, et que, par une récurrence immédiate,

$$\mathbb{E}\left(\frac{X_n}{m^n} \mid \mathcal{F}_0\right) = X_0 + \sum_{k=0}^n \frac{Z_k}{m^k}, \quad \text{pour tout } n \geq 1.$$

Le lemme 3.1.16 entraîne alors que $(\frac{X_n}{m^n})_n$ est une sous-martingale avec des espérances bornées, et donc elle converge p.s. vers une variable aléatoire finie p.s. (Pour la définition d'une sous-martingale et le théorème de convergence, voir MAP 432). \square

3.1.4 Les probabilités quasi-stationnaires

Nous venons de voir que dans les cas sous-critique ou critique, la chaîne de BGW s'éteint presque-sûrement. Nous verrons ultérieurement qu'il en est de même pour des processus décrivant certaines dynamiques de populations densité-dépendantes. Mais le temps nécessaire à l'extinction peut-être très long (par exemple dans le cas critique) et dans une échelle de temps qui n'est pas l'échelle de temps humaine. On peut alors dans certains cas observer une apparente stationnarité de la population avant l'extinction. Notre but dans ce paragraphe est de donner un sens mathématique à cette stabilité avant extinction. Mathématiquement, et par analogie avec la définition d'une probabilité invariante (ou stationnaire), on donnera la définition suivante.

Définition 3.1.17 Soit X un processus absorbé en 0 avec probabilité 1. On dira qu'une probabilité ν est une **probabilité quasi-stationnaire** pour le processus X si elle vérifie que pour tout $j \geq 1$ et pour tout n ,

$$\mathbb{P}_\nu(X_n = j | T > n) = \nu_j.$$

Nous allons justifier cette définition en étudiant les quantités $\mathbb{P}_i(X_n = j | T > n) = \mathbb{P}_i(X_n = j | X_n \neq 0)$ pour tout $j \in \mathbb{N}^*$, quand n est grand. Nous allons montrer que sous des hypothèses adéquates, les $\mathbb{P}_i(X_n = j | T > n)$ convergent, quand n tend vers l'infini et pour tout $i, j \geq 1$, vers des limites μ_j , bien-sûr positives, et telles que $\sum_j \mu_j = 1$. On vérifiera que la probabilité μ ainsi définie est une mesure quasi-stationnaire, au sens de la définition 3.1.17. On a donc une sorte de "théorème ergodique conditionnel". Ce phénomène n'est pas propre aux chaînes BGW, mais concerne une large classe de chaînes de Markov ayant un état absorbant.

Nous nous plaçons donc dans un cadre un peu plus général. On considère une chaîne de Markov $(X_n)_n$ de matrice de transition $P = (p_{ij})$, avec un unique état absorbant. On suppose que l'espace d'état est \mathbb{N} et que l'état absorbant est 0, donc notre hypothèse revient à dire que

$$p_{00} = 1, \quad \text{et } p_{ii} < 1 \text{ pour } i \geq 1. \quad (3.1.20)$$

Le temps d'absorption est comme précédemment $T = \inf\{n, X_n = 0\}$ et nous avons $X_n = 0$ si et seulement si $n \geq T$. On note toujours \mathbb{P}_i la probabilité pour laquelle $X_0 = i$.

Supposons aussi qu'il existe une fonction propre réelle φ définie sur \mathbb{N} associée à une valeur propre $\lambda > 0$ telles que

$$P\varphi = \lambda\varphi, \quad \varphi(0) = 0, \quad \inf_{i \geq 1} \varphi(i) > 0. \quad (3.1.21)$$

Remarquons que la première assertion s'écrit : pour tout $i \in \mathbb{N}$,

$$\sum_{j \in \mathbb{N}} p_{ij} \varphi(j) = \lambda \varphi(i).$$

Remarque : Cette hypothèse est vérifiée pour la chaîne de BGW en prenant pour fonction φ la fonction identité $\varphi(i) = i$ pour tout i , et $\lambda = m$.

Lemme 3.1.18 La formule $q_{ij} = \frac{p_{ij} \varphi(j)}{\lambda \varphi(i)}$ pour $i, j \geq 1$ définit une probabilité de transition $Q = (q_{ij})$ sur \mathbb{N}^* , dont la n ème puissance $Q^n = (q_{ij}^{(n)})$ est donnée par

$$q_{ij}^{(n)} = \frac{p_{ij}^{(n)} \varphi(j)}{\lambda^n \varphi(i)}. \quad (3.1.22)$$

Preuve. On a $q_{ij} \geq 0$, et puisque $\varphi(0) = 0$,

$$\sum_{j \geq 1} q_{ij} = \frac{1}{\lambda \varphi(i)} \sum_{j \geq 1} p_{ij} \varphi(j) = \frac{1}{\lambda \varphi(i)} \sum_{j \geq 0} p_{ij} \varphi(j) = \frac{1}{\lambda \varphi(i)} (P\varphi)(i) = \frac{\lambda \varphi(i)}{\lambda \varphi(i)} = 1,$$

où l'on a utilisé $P\varphi = \lambda\varphi$. Cela prouve la première assertion.

La relation (3.1.22) est vraie pour $n = 1$. Si on la suppose vraie pour $n - 1$, et en utilisant le fait que $p_{0j}^{(n-1)} = 0$ si $j \geq 1$, on a

$$q_{ij}^{(n)} = \sum_{k \geq 1} q_{ik} q_{kj}^{(n-1)} = \sum_{k \geq 1} \frac{p_{ik} \varphi(k)}{\lambda \varphi(i)} \frac{p_{kj}^{(n-1)} \varphi(j)}{\lambda^{n-1} \varphi(k)} \quad (3.1.23)$$

$$= \frac{\varphi(j)}{\lambda^n \varphi(i)} \sum_{k \geq 1} p_{ik} p_{kj}^{(n-1)} = \frac{\varphi(j)}{\lambda^n \varphi(i)} \sum_{k \geq 0} p_{ik} p_{kj}^{(n-1)} = \frac{p_{ij}^{(n)} \varphi(j)}{\lambda^n \varphi(i)}. \quad (3.1.24)$$

On a donc montré (3.1.22) par récurrence. \square

On déduit en particulier de (3.1.21) et (3.1.22) que $p_{ij}^{(n)} > 0 \Leftrightarrow q_{ij}^{(n)} > 0$ si $i, j \geq 1$. Par suite, si $C_0 = \{0\}$ et C_1, C_2, \dots désignent les classes de la chaîne (X_n) , il est immédiat que :

La période de l'état $i \geq 1$ est la même pour les chaînes de transitions P et Q , (3.1.25)

Les classes de la chaîne de transition Q sont C_1, C_2, \dots , (3.1.26)

Nous allons supposer ici que la chaîne de matrice de transition Q a une unique classe positive C apériodique. Nous allons voir dans la suite que c'est le cas du processus de BGW. Dans ce cas, cette chaîne admet une unique probabilité invariante π telle que $\pi_i > 0$ si $i \in C$, et pour tous i et j , $q_{ij}^{(n)} \rightarrow \pi_j$. (Voir MAP 432).

Théorème 3.1.19 a) Supposons (3.1.20) et (3.1.21) avec un $\lambda < 1$.

On a alors $\mathbb{P}_i(T < \infty) = 1$ pour tout i , et les états $i \geq 1$ sont transients pour la chaîne (X_n) .

b) Supposons en outre que la chaîne de transition Q admet une unique classe positive C , apériodique. Si $\pi = (\pi_i)_{i \geq 1}$ désigne l'unique probabilité invariante pour Q (donc $\pi_i > 0$ si et seulement si $i \in C$), on a pour tous $i, j \geq 1$:

$$\lim_{n \rightarrow \infty} \mathbb{P}_i(X_n = j | X_n \neq 0) = \mu_j := \frac{\pi_j / \varphi(j)}{\sum_{k \geq 1} \pi_k / \varphi(k)}, \quad (3.1.27)$$

et $\mu = (\mu_j)_{j \geq 1}$ est une probabilité, portée par C .

c) La probabilité μ est une probabilité quasi-stationnaire.

La probabilité μ s'appelle la **limite de Yaglom** de la chaîne (X_n) . On remarquera qu'elle ne dépend pas de la condition initiale i choisie.

Remarquons que l'événement conditionnant $\{X_n \neq 0\} = \{T \geq n\}$ a une probabilité qui tend vers 0 quand n tend vers l'infini. On conditionne donc par un événement de plus en plus rare, quand n augmente.

Preuve. (a) Si $a = \inf_{i \geq 1} \varphi(i)$, on a

$$a\mathbb{P}_i(T \geq n) = a \sum_{j \geq 1} p_{ij}^{(n)} \leq P^n \varphi(i) = \lambda^n \varphi(i),$$

qui tend vers 0 quand $n \rightarrow \infty$ puisque $\lambda < 1$. Comme $a > 0$, on en déduit la première assertion. De plus, si $i \geq 1$, $\mathbb{P}_i(T < \infty) = 1$ entraîne que la chaîne ne passe \mathbb{P}_i -p.s. qu'un nombre fini de fois en $i \geq 1$, puisque 0 est absorbant, et donc i est transient.

(b) Sous l'hypothèse supplémentaire, on a $q_{ij}^{(n)} \rightarrow \pi_j$. On a de plus que pour toute fonction bornée h , $\sum_{j \geq 1} q_{ij}^{(n)} h(j) \rightarrow \sum_{j \geq 1} \pi_j h(j)$. Mais pour $i, j \geq 1$, (3.1.22) implique

$$\mathbb{P}_i(X_n = j | X_n \neq 0) = \frac{p_{ij}^{(n)}}{\sum_{k \geq 1} p_{ik}^{(n)}} = \frac{q_{ij}^{(n)} / \varphi(j)}{\sum_{k \geq 1} q_{ik}^{(n)} / \varphi(k)}, \quad (3.1.28)$$

la fonction $1/\varphi$ étant bornée sur \mathbb{N}^* . La convergence (3.1.27) est alors immédiate.

La propriété $\sum_{j \geq 1} \mu_j = 1$ est évidente.

(c) Vérifions que μ est une probabilité quasi-stationnaire. Soit $j \geq 1$ et $n \in \mathbb{N}^*$. Alors

$$\begin{aligned} \mathbb{P}_\mu(X_n = j | X_n \neq 0) &= \frac{\mathbb{P}_\mu(X_n = j, X_n \neq 0)}{\mathbb{P}_\mu(X_n \neq 0)} \\ &= \frac{\sum_{k \geq 1} \mu_k \mathbb{P}_k(X_n = j, X_n \neq 0)}{\sum_{k \geq 1} \mu_k \mathbb{P}_k(X_n \neq 0)} = \frac{\sum_{k \geq 1} \mu_k \mathbb{P}_k(X_n = j)}{\sum_{k \geq 1} \mu_k \mathbb{P}_k(X_n \neq 0)} \\ &= \frac{\sum_{k \geq 1} \frac{\pi_k}{\varphi(k)} p_{kj}^{(n)}}{\sum_{k \geq 1} \frac{\pi_k}{\varphi(k)} p_{kj}^{(n)}} = \frac{\sum_{k \geq 1} \frac{\pi_k}{\varphi(k)} p_{kj}^{(n)}}{\sum_{i \geq 1} \left(\sum_{k \geq 1} \frac{\pi_k}{\varphi(k)} p_{ki}^{(n)} \right)} \\ &= \frac{\pi_j / \varphi(j)}{\sum_{i \geq 1} \pi_i / \varphi(i)} = \mu_j. \end{aligned}$$

Pour obtenir la dernière égalité, nous avons utilisé que π est une mesure invariante pour Q définie par (3.1.22). \square

Ce résultat nous donne le comportement limite (en loi) de X_n , sachant $X_n \neq 0$. Il est également intéressant de déterminer le comportement limite de la chaîne $(X_l)_{0 \leq l \leq N}$ à "horizon fini" N (arbitrairement grand, mais fixé), sachant que $X_n \neq 0$, lorsque $n \rightarrow \infty$. Cela revient à vouloir connaître la loi de $(X_l)_{0 \leq l \leq N}$ conditionnellement au fait que le processus ne s'éteindra jamais.

Théorème 3.1.20 *Plaçons-nous de nouveau sous les hypothèses du Théorème 3.1.19-(b). Soit $N \geq 1$ fixé. Pour tout $i \geq 1$, la chaîne $(X_l)_{1 \leq l \leq N}$ converge en loi, sous les probabilités conditionnelles sachant que $X_n \neq 0$ et lorsque $n \rightarrow \infty$, vers une chaîne de Markov à valeurs dans \mathbb{N}^* de matrice de transition Q .*

Cette nouvelle chaîne est souvent appelée le Q -processus associé à la chaîne initiale X . Remarquons qu'il est évident qu'elle doit vivre sur \mathbb{N}^* , puisque l'on a conditionné à la non-extinction. Par ailleurs, nous avons déjà vu que la probabilité invariante de cette chaîne est π . Remarquons, et cela peut paraître surprenant, que la limite de Yaglom μ et cette probabilité invariante π ne sont pas égales, contrairement à ce que l'intuition aurait pu laisser supposer.

Une manière équivalente d'énoncer le résultat du Théorème 3.1.20 consiste à dire que

$$\mathbb{P}_i(X_0 = i_0, X_1 = i_1, \dots, X_N = i_N | X_n \neq 0) \rightarrow \delta_{ii_0} q_{i_0 i_1} q_{i_1 i_2} \dots q_{i_{N-1} i_N} \quad (3.1.29)$$

pour tous $i_j \geq 1$, lorsque $n \rightarrow \infty$, et où δ_{ij} est le symbole de Kronecker.

Preuve. On va en fait montrer (3.1.29). Notons (pour i_0, \dots, i_N fixés) les membres de gauche et de droite de (3.1.29) par γ_n et γ respectivement. En utilisant (3.1.22), on obtient

$$\begin{aligned} \gamma_n &= \delta_{ii_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{N-1} i_N} \frac{\sum_{k \geq 1} p_{i_N k}^{(n-N)}}{\sum_{k \geq 1} p_{i_0 k}^{(n)}} \\ &= \delta_{ii_0} \frac{\lambda q_{i_0 i_1} \varphi(i_0)}{\varphi(i_1)} \frac{\lambda q_{i_1 i_2} \varphi(i_1)}{\varphi(i_2)} \dots \frac{\lambda q_{i_{N-1} i_N} \varphi(i_{N-1})}{\varphi(i_N)} \frac{\sum_{k \geq 1} \lambda^{n-N} q_{i_N k}^{(n-N)} \varphi(i_N) / \varphi(k)}{\sum_{k \geq 1} \lambda^n q_{i_N k}^{(n)} \varphi(i_0) / \varphi(k)} \\ &= \gamma \frac{\sum_{k \geq 1} q_{i_N k}^{(n-N)} / \varphi(k)}{\sum_{k \geq 1} q_{i_N k}^{(n)} / \varphi(k)}. \end{aligned}$$

On a vu dans la preuve précédente que $\sum_{k \geq 1} q_{i_N k}^{(n)} / \varphi(k) \rightarrow \sum_{k \geq 1} \pi_k / \varphi(k)$, de sorte que l'expression précédente tend vers γ , et on a le résultat. \square

Nous allons maintenant appliquer les résultats précédents dans le cas particulier de la chaîne de BGW, pour laquelle nous utilisons comme précédemment les notations m , m_2 pour les moments d'ordre 1 et 2 et la notation g pour la fonction génératrice de la loi de reproduction. Nous supposons (3.1.5), et C est la classe transiente décrite dans la Proposition 3.1.5. Nous allons étudier l'existence de mesures quasi-stationnaires dans le cas sous-critique où l'on sait que le processus s'éteint presque-sûrement.

Théorème 3.1.21 *Si $m < 1$ et $m_2 < \infty$, il existe une probabilité quasi-stationnaire $\mu = (\mu_i)_{i \geq 1}$. Pour $i, j \geq 1$, on a*

$$\mathbb{P}_i(X_n = j | X_n \neq 0) = \mathbb{P}_i(X_n = j | T > n) \rightarrow \mu_j,$$

et $\mu_j > 0$ si et seulement si $j \in C$.

De plus pour tout N fixé, la chaîne $(X_l)_{1 \leq l \leq N}$ converge en loi, sous la probabilité conditionnelle $\mathbb{P}_i(\cdot | X_n \neq 0)$ avec $i \geq 1$ et quand $n \rightarrow \infty$, vers une chaîne de Markov à valeurs dans \mathbb{N}^* partant de i et de transition $q_{ij} = \frac{j p_{ij}}{im}$.

Preuve. Nous avons vu ci-dessus que les hypothèses (3.1.21) sont satisfaites avec $\varphi(i) = i$ et $\lambda = m < 1$, et (3.1.20) est trivial. On associe à P les matrices de transition Q et Q^n définies par (3.1.22). D'après la preuve de la proposition 3.1.5, on sait que $p_{ij}^{(n)} = 0$ si $n \geq 1$ et $j \notin C \cup \{0\}$. Par suite, les singletons $\{j\}$ pour $j \notin C$ et $j \geq 1$ sont aussi des classes transientes pour la chaîne de transition Q . D'après (3.1.26), C est alors l'unique classe de la chaîne associée à Q qui peut être récurrente positive.

Posons $i_0 = \inf\{j \geq 1, q_j > 0\}$. (Rappelons que q_i est la probabilité d'avoir i descendants). Nous avons $p_{i_0 i_0} = i_0 q_{i_0}(q_0)^{i_0-1} > 0$, puisque la seule possibilité pour qu'un groupe de i_0 individus ait i_0 descendants est qu'un des individus ait i_0 descendants et les autres 0 (par hypothèse, on ne peut avoir moins de i_0 descendants, sauf 0). Nous en déduisons que la période de i_0 , donc de la classe C , est 1 pour la chaîne (X_n) , donc aussi pour celle associée à Q (par (3.1.25)).

Pour pouvoir appliquer les théorèmes 3.1.19 et 3.1.20, il nous suffit alors de montrer que $q_{1i}^{(n)}$ ne tend pas vers 0 pour tout $i \geq 1$. (Cela assurera que la classe est positive). Rappelons qu'ici,

$$q_{ij} = \frac{j p_{ij}}{im} ; q_{ij}^{(n)} = \frac{j p_{ij}^{(n)}}{im^n}$$

Nous allons montrer que $\Phi_n(s) = \sum_{i \geq 1} q_{1i}^{(n)} s^i$ converge vers une limite non nulle pour un $s \in]0, 1[$. Alors il existera $i \geq 1$ tel que $q_{1i}^{(n)}$ ne converge pas vers 0.

On a

$$\begin{aligned} m^{n+1} \Phi_{n+1}(s) &= \sum_{i \geq 1} p_{1i}^{(n+1)} i s^i = \mathbb{E}_1 (X_{n+1} s^{X_{n+1}}) \\ &= s g'_{n+1}(s) \\ &= s g'_n(g(s)) g'(s) \\ &= \frac{s g'(s)}{g(s)} \mathbb{E}_1 (X_n g(s)^{X_n}) = \frac{s g'(s)}{g(s)} m^n \Phi_n \circ g(s). \end{aligned}$$

Par suite, comme $\Phi_0(s) = s$, on a

$$\Phi_n(s) = s \prod_{k=0}^{n-1} \frac{g' \circ g_k(s)}{m}.$$

Notons que $0 < g'(s) < m$ pour tout $s \in [0, 1[$, donc $0 < \Phi_n(s) < 1$ si $0 < s < 1$. Il nous suffit alors de prouver que le produit infini de terme général $g' \circ g_k(s)/m$ est convergent

(ce qui signifie que $\Phi_n(s)$ converge vers une limite $\Phi(s) > 0$), et ceci est réalisé si la série de terme général $u_k := 1 - g' \circ g_k(s)/m$ est absolument convergente. En faisant un développement limité à l'ordre 1 de $g'(s)$ au voisinage de 1 (qui est possible car $m_2 < \infty$), on obtient que

$$1 - \frac{g'(s)}{m} = (1-s) \frac{g''(1)}{m} + (1-s)\varepsilon(1-s),$$

où la fonction ε tend vers 0 en 0. Il existe donc une constante a telle qu'on ait

$$0 \leq 1 - \frac{g'(s)}{m} \leq a(1-s).$$

On sait aussi que $0 \leq 1 - g_k(s) \leq m^k(1-s)$. Donc $0 \leq u_k \leq am^k(1-s)$, et on a le résultat puisque $m < 1$. \square

3.1.5 les chaînes densité-dépendantes

Nous allons considérer maintenant, de manière un peu superficielle, des chaînes BGW qui sont densité-dépendantes, et sans immigration. Bien sûr, ce sont des modèles beaucoup plus crédibles du point de vue de la biologie. Nous les développerons plus dans le cadre du temps continu. L'état 0 est toujours un état absorbant, et le temps d'extinction est encore $T = \inf\{n : X_n = 0\}$. En revanche, la "propriété de branchement" n'est plus valide.

Typiquement, une telle chaîne modélise l'évolution d'une population de type BGW, mais avec des "ressources limitées" qui diminuent les possibilités de reproduction lorsque l'effectif r de la population augmente. Ainsi, il est naturel de supposer les conditions suivantes sur les moyennes $m(r)$ des lois de reproduction $q(r)$ et les probabilités de mort sans descendance :

$$m(r+1) \leq m(r), \quad q_0(r) \leq q_0(r+1). \quad (3.1.30)$$

Par exemple, considérons un cadre d'interaction appelé "logistique", où le nombre total d'individus agit linéairement sur la probabilité $q_0(r)$. On peut modéliser cette situation ainsi :

$$q_0(r) = \frac{q_0 + Cr}{1 + Cr}; \quad q_i(r) = \frac{q_i}{1 + Cr},$$

où (q_i) est une probabilité sur \mathbb{N} et C une constante strictement positive. Ce modèle vérifie bien les propriétés (3.1.30).

On pourrait démontrer des résultats analogues à ceux des chaînes BGW standards, mais nous nous contenterons d'un résultat très partiel sur l'extinction. Rappelons que $m(r) < 1$ implique $q_0(r) > 0$, puisque $m(r) = \sum_{j \geq 1} j q_j(r) \geq \sum_{j \geq 1} q_j(r) = 1 - q_0(r)$.

Théorème 3.1.22 *Supposons que $q_0(r) > 0$ pour tout $r \geq 1$ et que $m(r) < \infty$ pour tout $r \geq 1$ et $m(r) \leq a$ sauf pour un nombre fini de r , avec un $a \in]0, 1[$. On a alors $\mathbb{P}_i(T < \infty) = 1$ (extinction p.s.).*

Notez que sous (3.1.30) les hypothèses de ce théorème se réduisent à $q_0(1) > 0$ et $\lim_{r \rightarrow \infty} m(r) < 1$.

Preuve. Comme $q_0(r) \geq 1 - m(r) \geq 1 - a$ si $m(r) \leq a$, on a $b = \inf_r q_0(r) > 0$, d'après les hypothèses.

Le même calcul que dans le cas densité-indépendant montre que $\mathbb{E}_i(X_{n+1} | \mathcal{F}_n) = X_n m(X_n)$ (avec $m(0) = 0$). Ainsi, la suite

$$Z_n = \begin{cases} X_0, & \text{si } n = 0 \\ X_n / \prod_{l=0}^{n-1} m(X_l) & \text{si } n \geq 1 \end{cases}$$

est une martingale positive, d'espérance $\mathbb{E}_i(Z_n) = i$. Elle converge alors \mathbb{P}_i -p.s. vers une limite finie Z . Par suite $M := \sup_n Z_n < \infty$ \mathbb{P}_i -p.s., et on a

$$X_n \leq M U_n, \quad \text{où } U_n = \prod_{l=0}^{n-1} m(X_l). \quad (3.1.31)$$

Soit $\alpha_N(i, j)$ le nombre de fois où X_n passe en $j \geq 1$ entre les instants 0 et N . Comme 0 est un état absorbant, il est égal au nombre de fois où $X_n = j$ et $X_{n+1} > 0$ pour n entre 0 et $N - 1$, plus éventuellement 1. Donc si $j \geq 1$,

$$\alpha_N(i, j) = \mathbb{E}_i\left(\sum_{n=0}^N 1_{\{X_n=j\}}\right) \leq 1 + \mathbb{E}_i\left(\sum_{n=0}^{N-1} 1_{\{X_n=j, X_{n+1}>0\}}\right).$$

Mais $p_{j0} = q_0(j)^j$, donc d'après la propriété de Markov,

$$\mathbb{P}_i(X_n = j, X_{n+1} > 0) = (1 - q_0(j)^j) \mathbb{E}_i(1_{\{X_n=j\}}) \leq (1 - b^j) \mathbb{E}_i(1_{\{X_n=j\}})$$

et on déduit que

$$\alpha_N(i, j) \leq 1 + (1 - b^j) \alpha_{N-1}(i, j).$$

En itérant cette relation et comme $\alpha_0(i, j) \leq 1$, on obtient

$$\alpha_N(i, j) \leq 1 + (1 - b^j) + \dots + (1 - b^j)^N \leq 1/b^j.$$

En faisant tendre N vers l'infini, on voit que

$$\mathbb{E}_i(N_j) < \infty,$$

où

$$N_j = \sum_n 1_{\{X_n=j\}}$$

est le nombre de visites de la chaîne en j .

Si C désigne l'ensemble (fini) des r tels que $m(r) > a$, on a

$$U_n \leq a^{\sum_{i=0}^{n-1} 1_{\{X_i \in C^c\}}} \prod_{j \in C} m(j)^{\alpha_{n-1}(i,j)}.$$

Mais comme chaque N_j est p.s. fini, on sait que le nombre de passages par C est fini p.s. Ainsi, $\sum_{i=0}^{n-1} 1_{\{X_i \in C^c\}}$ tend vers l'infini avec n . Par suite $U_n \rightarrow 0$ p.s., donc (3.1.31) entraîne que $X_n \rightarrow 0$ p.s. Comme X_n est à valeurs entières et que 0 est absorbant, cela donne le résultat. \square

3.2 Processus markoviens de saut en temps continu

Dans le chapitre précédent, nous avons considéré des processus de branchement indexés par un temps entier qui modélisait essentiellement un nombre de générations ou de périodes de temps dans un processus de reproduction cyclique (reproduction saisonnière, modèles de démographie dans lesquels on étudie des changements annuels). Maintenant, nous souhaitons modéliser une évolution de population composée d'individus qui naissent et meurent au cours du temps, mais en prenant en compte le temps physique qui évolue de manière continue : $t \in \mathbb{R}_+$.

La taille de la population est alors décrite par un processus aléatoire $t \rightarrow X_t$. A chaque temps de naissance, le processus s'accroît de 1, ou plus si on a une reproduction multiple, à chaque mort il décroît de 1. Les temps auxquels ont lieu les naissances et les morts sont aléatoires et il est nécessaire de connaître leur loi, de même que la loi de reproduction. Nous allons donc généraliser l'approche probabiliste des processus de branchement à temps discret. Toutefois, on peut aussi modéliser la dynamique continue en temps et raisonner comme les biologistes en ont l'habitude, en étudiant l'évolution au cours du temps de la probabilité d'être dans un état précis pour la population. Nous en donnons un exemple en Section 3.2.1. Cela conduit à étudier un système d'équations différentielles. Nous verrons dans la suite de ce chapitre le lien entre processus de saut en temps continu et équations de Kolmogorov.

3.2.1 Une approche intuitive

Donnons ici une description intuitive d'une population de cellules qui croît par division cellulaire. On suppose que pendant un intervalle de temps très court de longueur h , la probabilité pour qu'une cellule se divise soit bh , pour un paramètre $b > 0$ qui est le

même pour toutes les cellules et indépendant de la taille de la population et du temps. On suppose également que h est suffisamment petit pour que la probabilité d'avoir plus d'une naissance pendant $]t, t + h[$ soit négligeable. Si la population est de taille N au temps $t + h$, alors soit elle était de taille N au temps t et il n'y a pas eu de naissance pendant l'intervalle $]t, t + h[$, soit elle était de taille $N - 1$ au temps t et il y a eu une naissance pendant $]t, t + h[$. La probabilité d'avoir une naissance pendant $]t, t + h[$ à partir de la population de taille $N - 1$ est $bh(N - 1)$, et celle de ne pas avoir de naissance à partir d'une population de taille N est donc $1 - bhN$. Ainsi donc, si on appelle $P_N(t)$ la probabilité d'avoir une population de taille N au temps t , on aura

$$P_N(t + h) = P_N(t)(1 - bhN) + P_{N-1}(t)bh(N - 1).$$

En divisant les deux termes de cette équation par h et en faisant tendre h vers 0, on obtient donc

$$\frac{dP_N(t)}{dt} = -bNP_N(t) + b(N - 1)P_{N-1}(t).$$

Ainsi l'on voit sur cet exemple qu'au processus stochastique de division cellulaire est associé un système d'équations différentielles.

Nous allons maintenant formaliser toute cette approche.

3.2.2 Processus de Poisson

Dans ce paragraphe et le suivant, nous allons étudier des processus de Markov indicés par \mathbb{R}_+ , à valeurs dans un ensemble fini ou dénombrable qui sont constants entre des temps de sauts aléatoires. Ces processus changent donc d'état de manière aléatoire. On les appelle **processus markoviens de sauts**.

Définition

Nous allons introduire le prototype de ces processus, à savoir **le processus de Poisson**. Ce processus modélise des répartitions aléatoires de points sur \mathbb{R}_+ qui vont correspondre essentiellement pour nous dans la suite, à des instants de naissance ou de mort d'individus.

Définition 3.2.1 *Un processus ponctuel sur \mathbb{R}_+ se décrit par la donnée d'une suite croissante de temps aléatoires*

$$0 < T_1 < T_2 < \dots < T_n < \dots \quad p.s.$$

dans \mathbb{R}_+ , qui sont des variables aléatoires définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et qui vérifient

$$T_n \rightarrow +\infty \quad \text{presque-sûrement, quand } n \text{ tend vers l'infini.}$$

Les instants T_n modélisent des instants où se produisent des événements.

On pose

$$S_1 = T_1 ; S_2 = T_2 - T_1 ; \dots , S_n = T_n - T_{n-1}, \dots$$

Les variables aléatoires S_n modélisent les longueurs des intervalles ou **temps d'attente** entre deux événements successifs.

Définition 3.2.2 La **fonction aléatoire de comptage** $(N_t)_{t \geq 0}$ associée au processus ponctuel $\{T_n, n \in \mathbb{N}\}$ est définie par

$$N_t = \sup\{n, T_n \leq t\} = \sum_{j \geq 1} \mathbf{1}_{\{T_j \leq t\}}.$$

N_t est donc le nombre d'événements qui se sont produits avant l'instant t .

Notons que $N_0 = 0$ puisque $T_1 > 0$ et que pour tout t , $N_t < \infty$ puisque la suite (T_n) tend vers l'infini. Pour $0 \leq s < t$, $N_t - N_s$ est le nombre d'événements qui ont eu lieu pendant l'intervalle $]s, t]$.

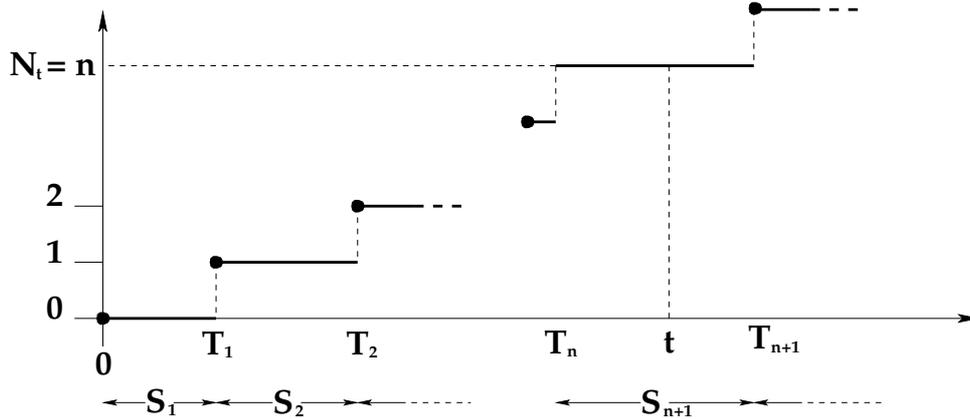


FIG. 3.3 – Trajectoire d'un processus ponctuel

Remarquons également que les trajectoires $t \rightarrow N_t(\omega)$ d'un tel processus sont continues à droite, par définition.

Remarque 3.2.3 les données du processus ponctuel et de la fonction aléatoire qui lui est associée sont en fait équivalentes. En effet, on a

$$\begin{aligned} \{N_t \geq n\} &= \{T_n \leq t\} \\ \{N_t = n\} &= \{T_n \leq t < T_{n+1}\} \\ \{N_t \geq n > N_s\} &= \{s < T_n \leq t\}. \end{aligned}$$

Définition 3.2.4 On dit que le processus ponctuel (T_n) ou $(N_t, t \geq 0)$ est un **processus de Poisson** si $(N_t, t \geq 0)$ est à **accroissements indépendants et stationnaires**, c'est à dire

1) Pour tous $t_0 < t_1 < \dots < t_n$ dans \mathbb{R}_+ , les accroissements $(N_{t_j} - N_{t_{j-1}}, 1 \leq j \leq n)$ sont des variables aléatoires indépendantes.

2) Pour $0 \leq s < t$, la loi de $N_t - N_s$ ne dépend de s et t que par la différence $t - s$. Elle est donc égale à la loi de N_{t-s} .

La propriété (2) s'appelle, comme dans le cas du mouvement brownien, la stationnarité des accroissements.

Le nom de processus de Poisson est justifié par la propriété suivante :

Proposition 3.2.5 Soit $(N_t, t \geq 0)$ un processus de Poisson. Alors il existe $\lambda > 0$ tel que pour tous $0 \leq s < t$, $N_t - N_s$ est une variable aléatoire de Poisson de paramètre $\lambda(t - s)$. On a donc

$$\mathbb{P}(N_t - N_s = k) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!}, \quad \forall k \in \mathbb{N}.$$

Définition 3.2.6 Ce paramètre λ est appelé **intensité** du processus de Poisson. Il est égal au nombre moyen d'événements qui se produisent pendant une unité de temps, puisque

$$\mathbb{E}(N_{t+1} - N_t) = \lambda.$$

On dira aussi que les événements se produisent **au taux** λ .

Preuve. Soit g_{t-s} la fonction génératrice de $N_t - N_s$. On a, pour $u \in [0, 1]$,

$$g_{t-s}(u) = \mathbb{E}(u^{N_t - N_s}) = \sum_{k \geq 0} \mathbb{P}(N_t - N_s = k) u^k.$$

Nous voulons montrer que g_{t-s} est la fonction génératrice d'une variable aléatoire de loi de Poisson de paramètre $\lambda(t - s)$, c'est-à-dire que

$$g_{t-s}(u) = \exp(-(\lambda(t-s)(1-u))).$$

Nous allons calculer $g_t(u)$. Par indépendance des accroissements, on a $g_t(u) = g_{t-s}(u)g_s(u)$, et donc plus généralement nous pouvons prouver que

$$g_t(u) = (g_1(u))^t.$$

(On le prouve pour les entiers puis pour les rationnels et on conclut en utilisant la décroissance de $t \rightarrow g_t(u)$).

Par ailleurs, comme $g_t(u) \geq \mathbb{P}(N_t = 0) = \mathbb{P}(T_1 > t)$, qui tend vers 1 quand $t \rightarrow 0$, on peut assurer que $g_1(u)$ est non nul, et comme $g_1(u) \leq 1$, il existe donc un $\lambda(u) > 0$ tel que

$$g_t(u) = e^{-\lambda(u)t}.$$

Nous souhaitons donc montrer que $\lambda(u)$ est de la forme $\lambda \times (1 - u)$, avec λ constante.

Nous avons

$$\lambda(u) = \lim_{t \downarrow 0} \frac{1}{t} (1 - g_t(u)) = \lim_{t \downarrow 0} \frac{1}{t} \sum_{k \geq 1} \mathbb{P}(N_t = k) (1 - u^k).$$

Puisque $u \geq 1$, nous en déduisons que pour tout t ,

$$0 \leq \frac{1}{t} (1 - g_t(u)) - \frac{1}{t} \mathbb{P}(N_t = 1) (1 - u) \leq \frac{1}{t} \mathbb{P}(N_t \geq 2).$$

Supposons que $\frac{1}{t} \mathbb{P}(N_t \geq 2)$ tende vers 0 quand $t \rightarrow 0$. Nous en déduisons alors que

$$\lambda(u) = \lim_{t \downarrow 0} \frac{1}{t} \mathbb{P}(N_t = 1) (1 - u),$$

et nous observons qu'ainsi $\lambda(u) = (1 - u)\lambda(0)$. Nous aurons alors

$$\lambda = \lambda(0).$$

Pour étudier le comportement asymptotique de $\mathbb{P}(N_2 \geq t)$, remarquons que

$$\cup_n \{N_{nt} = 0, N_{(n+1)t} \geq 2\} \subset \{T_2 < T_1 + t\},$$

et que par la propriété d'accroissements indépendants stationnaires,

$$\mathbb{P}(\cup_n \{N_{nt} = 0, N_{(n+1)t} \geq 2\}) = \sum_n \mathbb{P}(N_{nt} = 0) \mathbb{P}(N_t \geq 2) \leq \mathbb{P}(T_2 < T_1 + t).$$

Nous en déduisons que

$$\sum_n e^{-\lambda(0)nt} \mathbb{P}(N_t \geq 2) = (1 - e^{-\lambda(0)t})^{-1} \mathbb{P}(N_t \geq 2) \leq \mathbb{P}(T_2 < T_1 + t).$$

Mais, quand $t \downarrow 0$, cette dernière quantité vaut $\mathbb{P}(T_2 \leq T_1) = 0$. Comme pour t suffisamment petit, on a par ailleurs que $(\lambda(0)t)^{-1} \leq (1 - e^{-\lambda(0)t})^{-1}$, nous en déduisons finalement que $\frac{1}{t} \mathbb{P}(N_t \geq 2)$ tend vers 0 quand $t \rightarrow 0$. \square

Notons que la Proposition 3.2.5 et la propriété d'indépendance des accroissements permet d'obtenir la loi de tout vecteur $(N_{t_1}, \dots, N_{t_d})$, pour $t_1 < \dots < t_d$.

Remarque 3.2.7 Nous pouvons donner une interprétation intuitive de ce résultat. Il résulte de la preuve ci-dessus que

$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t = 0) &= 1 - \lambda h + o(h) \\ \mathbb{P}(N_{t+h} - N_t = 1) &= \lambda h + o(h) \\ \mathbb{P}(N_{t+h} - N_t \geq 2) &= o(h).\end{aligned}$$

Donc à une probabilité petite devant h près, $N_{t+h} - N_t$ est une variable aléatoire de Bernoulli prenant la valeur 0 avec probabilité $1 - \lambda h$ et la valeur 1 avec probabilité λh . Cette propriété jointe à l'indépendance des accroissements et à la formule

$$N_{t+s} - N_t = \sum_{j=1}^n (N_{t+jh} - N_{t+(j-1)h}), \quad \text{avec } h = \frac{s}{n},$$

entraîne que $N_{t+s} - N_t$ suit approximativement une loi binomiale de paramètre $(n, \lambda s/n)$. Nous savons que quand n tend vers l'infini, cette loi tend vers une loi de Poisson de paramètre λs . (Voir Polycopié de MAP 311).

Nous pouvons déduire de la Proposition 3.2.5 la loi du premier temps de saut du processus (N_t) .

Corollaire 3.2.8 *La loi du premier temps de saut T_1 est une loi exponentielle de paramètre λ . De même, pour tout $s > 0$, la loi du premier événement après s , soit $T_{N_s+1} - s$, est une loi exponentielle de paramètre λ .*

Preuve. Pour $t > 0$, on a

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}.$$

De même,

$$\mathbb{P}(T_{N_s+1} - s > t) = \mathbb{P}(N_{s+t} - N_s = 0) = \mathbb{P}(N_t = 0).$$

□

Propriété de Markov

Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ . Pour tout $s > 0$, introduisons le processus $(N_t^s, t \geq 0)$ défini par

$$N_t^s = N_{t+s} - N_s.$$

Ce processus compte le nombre d'événements sur l'intervalle $]s, t]$. On vérifie immédiatement que c'est également un processus de Poisson d'intensité λ , indépendant de $(N_u, u \leq s)$. En particulier, cela entraîne que le futur (N_{t+s}) du processus après le temps s ne dépend du passé $(N_u, u \leq s)$ que par l'intermédiaire du présent N_s . Nous avons

ainsi obtenu une version en temps continu de la propriété de Markov. Nous reviendrons sur cette propriété dans le paragraphe suivant.

Nous allons généraliser cette propriété à certains instants S aléatoires.

Remarquons que la donnée de $(N_s, s \leq t)$ est équivalente à celle de $(N_t, T_1, T_2, \dots, T_{N_t})$. Nous allons appeler \mathcal{F}_t^N la tribu engendrée par ces variables, qui décrit donc toute l'information donnée par le processus (N_t) jusqu'au temps t .

Définition 3.2.9 *Etant donné un processus de Poisson $(N_t, t \geq 0)$, on appelle temps d'arrêt (pour le processus de Poisson) une variable aléatoire S à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ telle que pour tout $t \in \mathbb{R}_+$,*

$$\{S \leq t\} \in \mathcal{F}_t^N.$$

Toute variable aléatoire constante est clairement un temps d'arrêt. Pour tout n , le temps de saut T_n l'est également. En revanche, T_{N_s} ne l'est pas car si $t < s$,

$$\{T_{N_s} \leq t\} = \{N_s - N_t = 0\} \notin \mathcal{F}_t^N.$$

Proposition 3.2.10 *Soit $(N_t, t \geq 0)$ un processus de Poisson de paramètre λ , et S un temps d'arrêt pour (N_t) . Sur l'événement $\{S < \infty\}$, on pose pour $t \geq 0$*

$$N_t^S = N_{S+t} - N_S.$$

Conditionnellement à $\{S < \infty\}$, le processus $(N_t^S, t \geq 0)$ est un processus de Poisson d'intensité λ , indépendant de la tribu engendrée par la trajectoire de N jusqu'à S .

Preuve. Nous savons déjà que le résultat est vrai si S est constant. Supposons que S prenne ses valeurs dans une suite croissante de réels positifs $(s_j)_j$. Comme S est un temps d'arrêt,

$$\{S = s_j\} = \{S \leq s_j\} \setminus \{S \leq s_{j-1}\} \in \mathcal{F}_{s_j}^N.$$

Soit $0 = t_0 < t_1 < t_2 < \dots < t_k$, $n_1, n_2, \dots, n_k \in \mathbb{N}$ et A un événement de la tribu engendrée par la trajectoire de N jusqu'à S . On a $A \cap \{S \leq t\} \in \mathcal{F}_t^N$. Alors

$$\begin{aligned} & \mathbb{P}\left(A \cap_{i=1}^k \{N_{t_i}^S - N_{t_{i-1}}^S = n_i\}\right) \\ &= \sum_j \mathbb{P}\left(\{S = s_j\} \cap A \cap_{i=1}^k \{N_{s_j+t_i} - N_{s_j} - N_{s_j+t_{i-1}} + N_{s_j} = n_i\}\right) \\ &= \sum_j \mathbb{P}\left(\{S = s_j\} \cap A\right) \prod_{i=1}^k \mathbb{P}\left(N_{s_j+t_i} - N_{s_j+t_{i-1}} = n_i\right) \\ &= \mathbb{P}(A) \prod_{i=1}^k \mathbb{P}\left(N_{t_i - t_{i-1}} = n_i\right). \end{aligned} \tag{3.2.32}$$

Dans cette preuve nous avons utilisé le fait que (N_t) est à accroissements indépendants et stationnaires.

Le résultat est donc établi si la suite des valeurs de S est discrète. Supposons maintenant que S est un temps d'arrêt quelconque. Nous introduisons la suite $(S_n)_n$ définie par

$$S_n = \sum_{k \in \mathbb{N}} k 2^{-n} \mathbf{1}_{\{(k-1)2^{-n} < S \leq k 2^{-n}\}}.$$

Il est facile de voir que les S_n sont des temps d'arrêt et ainsi, S peut être approché par une suite de temps d'arrêt de la forme précédente. L'égalité (3.2.32) est vraie pour chaque S_n et on peut facilement justifier un passage à la limite, du fait de la continuité à droite des trajectoires de (N_t) . Cela conclut la preuve. \square

Nous en déduisons le résultat fondamental suivant.

Proposition 3.2.11 *Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ et $(T_n)_n$ ses instants de saut. Notons comme précédemment par $(S_n)_n$ la suite des temps d'attente entre les sauts. Alors les variables S_i sont indépendantes, de loi exponentielle de paramètre λ .*

Preuve. On sait déjà que $S_1 = T_1$ vérifie cette propriété. On applique alors la Proposition 3.2.10 avec $S = T_n$. Alors $S_{n+1} = T_{n+1} - T_n$ est le premier instant de saut du processus de Poisson $(N_{T_n+t} - N_{T_n}, t \geq 0)$ d'intensité λ et indépendant de T_1, \dots, T_n , donc aussi de S_1, \dots, S_n . \square

On a également la réciproque suivante.

Proposition 3.2.12 *Soit $(S_n)_n$ une suite de variables aléatoires exponentielles de paramètre λ et indépendantes. On pose pour tout $n \geq 1$ et tout $t > 0$*

$$\begin{aligned} T_n &= S_1 + \dots + S_n, \\ N_t &= \sup\{n, T_n \leq t\} = \sum_n \mathbf{1}_{\{T_n \leq t\}} = \text{Card}\{n, T_n \leq t\}. \end{aligned}$$

Alors $(N_t, t \geq 0)$ est un processus de Poisson d'intensité λ .

Cette proposition est immédiate mais fournit une preuve constructive de l'existence d'un processus de Poisson.

Comportement asymptotique

Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ . On a alors

$$\mathbb{E}(N_t) = \lambda t, \quad \text{Var}(N_t) = \lambda t.$$

Ainsi,

$$\mathbb{E}(t^{-1} N_t) = \lambda, \quad \text{Var}(t^{-1} N_t) = \frac{\lambda}{t},$$

donc $\frac{N_t}{t}$ converge en moyenne quadratique vers λ , quand t tend vers l'infini. En fait on a également une version forte de cette loi des grands nombres.

Proposition 3.2.13 *Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ . Alors $\frac{N_t}{t}$ converge presque-sûrement vers λ , quand t tend vers l'infini.*

Preuve. Remarquons tout d'abord que $N_n = \sum_{i=1}^n (N_i - N_{i-1})$ est la somme de variables aléatoires indépendantes de même loi de Poisson de paramètre λ . Il résulte alors de la loi forte des grands nombres que $\frac{N_n}{n}$ converge p.s. vers λ , quand n tend vers l'infini. On écrit alors

$$\frac{N_t}{t} = \frac{N_{[t]}}{[t]} \frac{[t]}{t} + \frac{N_t - N_{[t]}}{t}.$$

Il nous suffit alors de montrer que $\sup_{n \leq t < n+1} \frac{N_t - N_n}{n}$ tend vers 0 quand n tend vers l'infini. Posons

$$\xi_n = \sup_{n \leq t < n+1} (N_t - N_n) = N_{n+1} - N_n.$$

Les ξ_n sont indépendantes et de même loi intégrable. Donc la loi des grands nombres entraîne que $\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow \lambda$ p.s. et donc que $\frac{\xi_n}{n}$ converge presque-sûrement vers 0. \square

On a également un théorème de la limite centrale.

Théorème 3.2.14 *Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ . Alors*

$$\frac{N_t - \lambda t}{\sqrt{\lambda t}} \rightarrow Z \quad \text{en loi, quand } t \rightarrow \infty,$$

où Z est une variable aléatoire de loi normale centrée réduite.

Preuve. Nous allons raisonner comme dans la preuve précédente. Par le théorème de la limite centrale, nous savons que

$$\frac{N_n - \lambda n}{\sqrt{\lambda n}} \rightarrow Z \quad \text{en loi, quand } n \rightarrow \infty,$$

où Z est une variable aléatoire de loi normale centrée réduite. De plus, $\frac{N_t - N_{[t]}}{\sqrt{\lambda t}} \leq \frac{\xi_{[t]}}{\sqrt{\lambda t}}$. Or,

$$\mathbb{P}\left(\frac{\xi_n}{\sqrt{n}} > \varepsilon\right) = \mathbb{P}(\xi_n > \varepsilon\sqrt{n}) = \mathbb{P}(\xi_1 > \varepsilon\sqrt{n}),$$

qui tend vers 0 quand n tend vers l'infini. Donc $\frac{N_t - N_{[t]}}{\sqrt{\lambda t}} \leq \frac{\xi_{[t]}}{\sqrt{\lambda t}}$ tend en probabilité vers 0. Finalement nous avons

$$\begin{aligned} \frac{N_t - \lambda t}{\sqrt{\lambda t}} &= \frac{N_{[t]} - \lambda[t]}{\sqrt{\lambda[t]}} \times \sqrt{\frac{[t]}{t}} \\ &\quad + \frac{N_t - N_{[t]}}{\sqrt{\lambda t}} + \sqrt{\lambda} \frac{[t] - t}{\sqrt{t}}. \end{aligned}$$

On utilise alors le même argument qu'en Section 2.1.1 pour conclure, sachant que $N_t - N_{[t]}$ et $N_{[t]}$ sont indépendantes. \square

Nous pouvons en fait établir également un théorème de la limite centrale fonctionnel, qui montre la convergence de la fonction aléatoire

$$t \rightarrow Y_{ut} = \frac{N_{ut} - \lambda t u}{\sqrt{\lambda u}}$$

vers le mouvement brownien, quand u tend vers l'infini. La preuve s'inspire de celle déjà vue en Section 2.1.1. En effet, nous avons montré que pour t fixé, Y_{ut} converge en loi vers une variable aléatoire B_t centrée, de variance t .

De même nous pouvons étudier le comportement en loi de $(Y_{ut_1}, \dots, Y_{ut_2} - Y_{ut_1}, Y_{ut_k} - Y_{ut_{k-1}})$, pour $t_1 < \dots < t_k$, et montrer que ce vecteur converge vers $(Z_{t_1}, Z_{t_2 - t_1}, \dots, Z_{t_k - t_{k-1}})$, où les variables aléatoires $Z_{t_i - t_{i-1}}$ sont indépendantes et suivent des lois normales centrées de variances respectives $t_i - t_{i-1}$. Remarquons par ailleurs que l'amplitude des sauts de Y_{ut} est $\frac{1}{\sqrt{\lambda u}}$ qui tend vers 0 quand u tend vers l'infini. On a donc finalement montré que le processus $(Y_{ut}, t \geq 0)$ converge au sens des marginales fini-dimensionnelles, vers un processus $(B_t, t \geq 0)$ qui est à accroissements indépendants et stationnaires, à trajectoires continues et tel que pour chaque $t > 0$ la variable aléatoire B_t est une variable normale centrée et de variance t .

3.2.3 Processus markoviens de saut

Nous allons ici étudier les processus de Markov en temps continu, à valeurs dans un espace fini ou dénombrable. Nous verrons que ces processus peuvent se décomposer à l'aide d'un processus de Poisson et d'une chaîne de Markov (à temps discret) appelée chaîne incluse.

Définition

On veut étudier des processus dont la trajectoire est continue à droite et limitée à gauche, constante entre des instants de saut qui sont isolés. Une grande différence avec le processus de Poisson est que les amplitudes des sauts sont aléatoires.

On notera Z_n la position du processus juste avant le saut T_{n+1} .

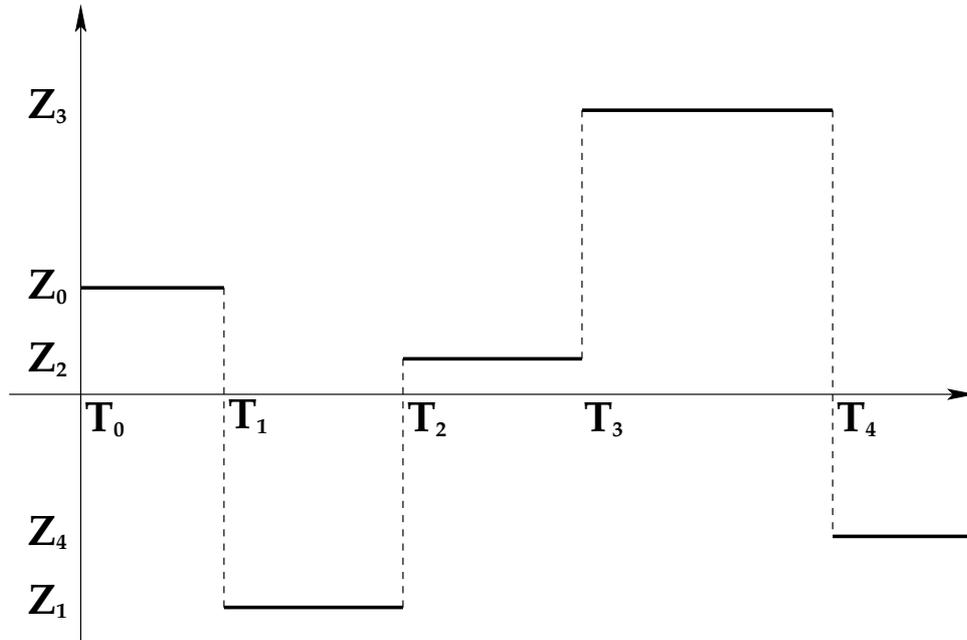


FIG. 3.4 – Trajectoire d'un processus markovien de saut

Une trajectoire type sera de la forme :

La donnée du processus $(X_t, t \geq 0)$ est donc équivalente à la donnée de la suite $(T_n, Z_n), n \geq 0$, avec $T_0 = 0$.

On suppose que les instants de saut $T_n \in \mathbb{R} \cup \{+\infty\}$ forment une suite croissante et que

$$T_n(\omega) < T_{n+1}(\omega) \text{ si } T_n(\omega) < \infty.$$

On suppose également qu'il n'y a pas d'explosion, c'est-à-dire que les instants de saut ne s'accumulent pas en un temps fini :

$$\lim_{n \rightarrow \infty} T_n(\omega) = +\infty. \quad (3.2.33)$$

Définition 3.2.15 On dira qu'un état x est **un état absorbant** si

$$X_{T_n}(\omega) = x \Rightarrow T_{n+1}(\omega) = +\infty.$$

Dans les études de dynamiques des populations, l'état 0 sera un état absorbant, sauf s'il y a de l'immigration.

Définition 3.2.16 On appelle **processus de saut** un processus de la forme

$$X_t(\omega) = \sum_{n \geq 0, T_n(\omega) < \infty} Z_n(\omega) \mathbf{1}_{[T_n(\omega), T_{n+1}(\omega)[}(t),$$

où les Z_n prennent leurs valeurs dans un ensemble dénombrable E et où les T_n vérifient les propriétés décrites ci-dessus.

On va étudier de tels processus. On les suppose markoviens, dans le sens suivant.

Définition 3.2.17 *Un processus de saut $(X_t, t \geq 0)$ à valeurs dans E est dit **processus markovien de saut** (ou chaîne de Markov en temps continu), 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 $0 \leq t_0 < t_1 < \dots < t_n < s$, $x_0, x_1, \dots, x_n, x, y \in E$, on a*

$$\mathbb{P}(X_t = y \mid X_{t_0} = x_0, X_{t_1} = x_1, \dots, X_{t_n} = x_n, X_s = x) = \mathbb{P}(X_t = y \mid X_s = x).$$

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

Dans le cas d'un processus homogène, on utilisera la notation

$$\mathbb{P}(X_t = y \mid X_s = x) = P_{xy}(t - s),$$

et pour tout $t > 0$, $P(t) = (P_{xy}(t))_{x,y \in E}$ est une matrice markovienne sur $E \times E$, (matrice de taille possiblement infinie dont les sommes des éléments de chaque ligne valent 1), appelée **matrice de transition de X_t sur E** . On notera $\mu(t)$ la loi de X_t et $\mu(0) = \mu$ désigne la loi initiale du processus.

Dans la suite, nous allons développer le cas homogène. Puisque E est dénombrable, on identifiera une fonction $g : E \rightarrow \mathbb{R}$ avec le vecteur $(g_y)_{y \in E}$, où $g_y = g(y)$, et une mesure μ sur E avec le vecteur (μ_x) où $\mu_x = \mu(\{x\})$.

Proposition 3.2.18 *Soit $(X_t)_t$ un processus markovien de saut homogène, de loi initiale μ et de matrice de transition $P(t), t > 0$. Pour tout $n \in \mathbb{N}$, pour tous $0 < t_1 < \dots < t_n$, la loi du vecteur aléatoire $(X_0, X_{t_1}, \dots, X_{t_n})$ est donnée par : pour tous $x_0, x_1, \dots, x_n \in E$,*

$$\mathbb{P}(X_0 = x_0, X_{t_1} = x_1, \dots, X_{t_n} = x_n) = \mu_{x_0} P_{x_0 x_1}(t_1) P_{x_1 x_2}(t_2 - t_1) \times \dots \times P_{x_{n-1} x_n}(t_n - t_{n-1}). \quad (3.2.34)$$

En particulier, pour tout $t > 0$,

$$\mu(t) = \mu P(t),$$

au sens où pour tout $y \in E$,

$$\mu_y(t) = \sum_{x \in E} \mu_x P_{xy}(t).$$

De plus, pour toute fonction positive ou bornée $g : E \rightarrow \mathbb{R}$, on a

$$\mathbb{E}(g(X_t) \mid X_0 = x) = (P(t)g)_x = \sum_{y \in E} P_{xy}(t) g_y.$$

En outre, les matrices $(P(t), t > 0)$ vérifient la relation de semi-groupe, appelée **équation de Chapman-Kolmogorov** :

$$P(t+s) = P(t)P(s), \quad (3.2.35)$$

où le produit (qui commute donc) est compris au sens matriciel.

Preuve. Soit $x, z \in E$. On a

$$\begin{aligned} P_{xz}(t+s) &= \mathbb{P}(X_{t+s} = z | X_0 = x) = \sum_y P(X_t = y, X_{t+s} = z | X_0 = x) \\ &= \sum_y \mathbb{P}(X_{t+s} = z | X_t = y) \mathbb{P}(X_t = y | X_0 = x) = \sum_y P_{xy}(t) P_{yz}(s), \end{aligned}$$

d'où $P(t+s) = P(t)P(s)$. □

Exemple 3.2.19 Un processus de Poisson $(N_t, t \geq 0)$ d'intensité λ est markovien, de matrice de transition

$$P_{xy}(t) = e^{-\lambda t} \frac{(\lambda t)^{y-x}}{(y-x)!}, \quad \text{si } y \geq x,$$

et $P_{xy}(t) = 0$ sinon.

Exemple 3.2.20 Soit $(N_t, t \geq 0)$ un processus de Poisson d'intensité λ et d'instantanés de sauts $(T_n)_n$. On se donne par ailleurs une chaîne de Markov $(Z_n)_n$ à valeurs dans E , indépendante de $(N_t)_t$ et de matrice de transition M_{xy} . Alors

$$X_t = \sum_{n=0}^{\infty} Z_n \mathbf{1}_{[T_n, T_{n+1}[}(t)$$

est un processus markovien de saut de matrice de transition

$$P_{xy}(t) = \sum_n \mathbb{P}(T_n \leq t < T_{n+1}, Z_n = y | Z_0 = x) = \sum_n \mathbb{P}(N_t = n) \mathbb{P}(Z_n = y | Z_0 = x),$$

par indépendance. D'où

$$P_{xy}(t) = e^{-\lambda t} \sum_n \frac{(\lambda t)^n}{n!} M_{xy}^{(n)}.$$

Propriété de Markov forte

On note \mathcal{F}_t^X la tribu engendrée par le processus X jusqu'au temps t , et par S un \mathcal{F}_t^X -temps d'arrêt. On a la propriété de Markov forte suivante, dont la preuve peut s'adapter de celle donnée pour le processus de Poisson.

Théorème 3.2.21 Soit S un temps d'arrêt pour le processus markovien de saut X_t . Alors, conditionnellement à $\{S < \infty\}$ et à $\{X_S = x\}$, le processus $(X_{S+t}, t \geq 0)$ est indépendant de la tribu engendrée par X jusqu'au temps S et sa loi est celle de $(X_t, t \geq 0)$ sachant que $X_0 = x$.

Le générateur infinitésimal

La propriété de semi-groupe (3.2.35) entraîne que $P(t)$ est connu pour tout t dès qu'il est connu pour t petit. En fait, nous allons montrer qu'il suffit de connaître sa dérivée à droite en 0.

Théorème 3.2.22 *Soit $P(t), t > 0$, le semi-groupe de matrices de transition d'un processus markovien de saut $(X_t)_t$. Il existe une matrice $(Q_{xy}, x, y \in E)$, appelée **générateur infinitésimal** du semi-groupe $P(t)$ ou du processus de Markov (X_t) , qui vérifie*

$$\begin{aligned} Q_{xy} &\geq 0 \text{ pour } x \neq y; \\ Q_{xx} &= - \sum_{y \in E \setminus \{x\}} Q_{xy} \leq 0, \end{aligned}$$

cette dernière inégalité étant stricte sauf si l'état x est absorbant.

Lorsque $h \downarrow 0$,

$$\begin{aligned} P_{xy}(h) &= hQ_{xy} + o(h) \text{ pour } x \neq y, \\ P_{xx}(h) &= 1 + hQ_{xx} + o(h). \end{aligned}$$

En outre conditionnellement à $X_0 = x$, l'instant T_1 de premier saut et la position $Z_1 = X_{T_1}$ après le premier saut sont indépendants, avec T_1 de loi exponentielle de paramètre $q_x := -Q_{xx}$ et Z_1 de loi sur E donnée par $\left(\frac{Q_{xy}}{q_x}, y \neq x\right)$.

Définition 3.2.23 *On appelle q_x le **taux de saut** du processus issu de x . On a donc*

$$\mathbb{P}(T_1 > t | X_0 = x) = e^{-q_x t}.$$

Par analogie avec les idées intuitives de l'introduction, le nombre Q_{xy} sera appelé **taux de transition de x vers y** .

Grâce à la propriété de Markov forte, nous en déduisons le résultat suivant.

Corollaire 3.2.24 *Pour tout $n \in \mathbb{N}^*$, conditionnellement à la tribu engendrée par le processus jusqu'au temps T_{n-1} , la variable aléatoire $T_n - T_{n-1}$ est indépendante de $Z_n = X_{T_n}$, la loi conditionnelle de $T_n - T_{n-1}$ est une loi exponentielle de paramètre $q_{Z_{n-1}}$, et la loi conditionnelle de $Z_n - Z_{n-1}$ est donnée par $\left(\frac{Q_{Z_{n-1}y}}{q_{Z_{n-1}}}, y \in E\right)$.*

Preuve. du théorème. Nous voulons calculer $\mathbb{P}(T_1 > t | X_0 = x)$. Considérons $n \in \mathbb{N}$ et $h > 0$ et supposons que $h \rightarrow 0$ et que $n \rightarrow \infty$ de telle sorte que $nh \uparrow t$. Nous allons utiliser la propriété de semi-groupe. Pour cela, remarquons tout d'abord que

$$\{T_1 > nh\} \subset \{X_0 = X_h = \dots = X_{nh}\} \subset \{T_1 > nh\} \cup \{T_2 - T_1 \leq h\}.$$

Comme $\mathbb{P}(T_2 - T_1 \leq h) \rightarrow 0$ quand $h \rightarrow 0$, on a que

$$\begin{aligned} \mathbb{P}(T_1 > t | X_0 = x) &= \lim_{h \rightarrow 0, nh \rightarrow t} \mathbb{P}(X_0 = X_1 = \dots = X_{nh} | X_0 = x) \\ &= \lim_{h \rightarrow 0, nh \rightarrow t} (P_{xx}(h))^n = \lim e^{n \ln P_{xx}(h)}, \text{ par propriété de Markov.} \end{aligned}$$

Comme $P_{xx}(h)$ tend vers 1 quand $h \rightarrow 0$, on a $\ln P_{xx}(h) \sim P_{xx}(h) - 1$ au voisinage de 0. Ainsi, comme de plus n est d'ordre $\frac{t}{h}$, nous pouvons déduire de l'existence de la limite précédente qu'il existe $q_x \in [0, +\infty]$ tel que

$$\lim_{h \rightarrow 0} \frac{1}{h} (1 - P_{xx}(h)) = q_x.$$

Ainsi,

$$\mathbb{P}(T_1 > t | X_0 = x) = e^{-q_x t}.$$

Cette dernière propriété entraîne que $q_x < \infty$ et que $q_x = 0$ si et seulement si x est un état absorbant. On pose alors $Q_{xx} = -q_x$.

La démonstration de l'existence d'une limite à $\frac{P_{xy}(h)}{h}$, pour $x \neq y$, se fait de manière analogue. On a

$$\begin{aligned} &\{T_1 \leq t, Z_0 = x, Z_1 = y\} \\ &= \lim_{h \rightarrow 0, nh \rightarrow t} \cup_{0 \leq m \leq n} \{X_0 = X_h = \dots = X_{(m-1)h} = x, X_{mh} = y\}. \end{aligned}$$

D'où

$$\begin{aligned} \mathbb{P}(T_1 \leq t, Z_1 = y | X_0 = x) &= \lim \sum_{m=0}^{n-1} (P_{xx}(h))^m P_{xy}(h) = \lim \frac{1 - (P_{xx}(h))^n}{1 - P_{xx}(h)} P_{xy}(h) \\ &= \lim (1 - (P_{xx}(h))^n) \frac{h}{1 - P_{xx}(h)} \frac{1}{h} P_{xy}(h) \\ &= \frac{1 - e^{-q_x t}}{q_x} \lim \frac{1}{h} P_{xy}(h). \end{aligned} \tag{3.2.36}$$

Ainsi, $Q_{xy} = \lim \frac{1}{h} P_{xy}(h)$ existe pour $x \neq y$ et

$$\mathbb{P}(T_1 \leq t, Z_1 = y | X_0 = x) = (1 - e^{-q_x t}) \frac{Q_{xy}}{q_x},$$

d'où

$$\mathbb{P}(T_1 \leq t, Z_1 = y | X_0 = x) = \mathbb{P}(T_1 \leq t | X_0 = x) \mathbb{P}(Z_1 = y | X_0 = x),$$

et on a donc indépendance de T_1 et Z_1 conditionnellement à $X_0 = x$. De plus,

$$\mathbb{P}(Z_1 = y | X_0 = x) = \frac{Q_{xy}}{q_x}. \tag{3.2.37}$$

Enfin, on déduit de (3.2.36) que pour h suffisamment petit,

$$\frac{1}{h}P_{xy}(h) \leq 2 \frac{q_x}{1 - e^{-q_x t}} \mathbb{P}(T_1 \leq t, Z_1 = y | X_0 = x).$$

Nous pouvons alors en déduire (théorème de convergence dominée) que

$$\lim_h \sum_y \frac{1}{h}P_{xy}(h) = \sum_y \lim_h \frac{1}{h}P_{xy}(h).$$

De cela et du fait que $\sum_y P_{xy}(h) = 1$ se déduit facilement que $\sum_y Q_{xy} = 0$. \square

On en déduit les équations de Kolmogorov, qui vont nous permettre de retrouver les matrices de transitions à partir de la matrice de taux. Soit I la matrice identité sur E .

Théorème 3.2.25 *Sous l'hypothèse (3.2.33),*

1) $(P(t), t \geq 0)$ est l'unique solution de l'équation de Kolmogorov rétrograde

$$\frac{dP}{dt}(t) = QP(t), \quad \text{pour } t > 0; P(0) = I, \quad (3.2.38)$$

c'est-à-dire

$$\frac{dP_{xy}}{dt}(t) = \sum_{z \in E} Q_{xz} P_{zy}(t). \quad (3.2.39)$$

Pour tout x de E et toute fonction g , $u(t, x) = \mathbb{E}(g(X_t) | X_0 = x)$ est solution de

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= \sum_{z \in E} Q_{xz} u(t, z), \quad t > 0, x \in E, \\ u(0, x) &= g(x), \quad x \in E. \end{aligned}$$

2) $(P(t), t \geq 0)$ est l'unique solution de l'équation de Kolmogorov progressive

$$\frac{dP}{dt}(t) = P(t)Q, \quad \text{pour } t > 0; P(0) = I, \quad (3.2.40)$$

c'est-à-dire

$$\frac{dP_{xy}}{dt}(t) = \sum_{z \in E} P_{xz}(t) Q_{zy}. \quad (3.2.41)$$

En outre, la famille des lois marginales $\mu(t)$ des X_t satisfait l'équation de Fokker-Planck

$$\frac{\partial \mu_y(t)}{\partial t} = \sum_{z \in E} \mu_z(t) Q_{zy}, \quad t > 0, y \in E.$$

Preuve. La preuve est très simple dans le cas où E est fini et plus technique sinon, mais les idées sont les suivantes. Pour établir l'équation de Kolmogorov rétrograde, il suffit de dériver $P(t+h)$ en $h=0$ en utilisant la propriété de semi-groupe sous la forme

$$P(t+h) = P(h)P(t).$$

L'équation pour u s'en déduit en multipliant à droite par le vecteur colonne (g_y) . L'équation progressive s'obtient de la même manière, mais en écrivant

$$P(t+h) = P(t)P(h).$$

L'équation de Fokker-Planck s'en déduit alors immédiatement en multipliant à gauche par le vecteur ligne $(\mu_x)(0)$. \square

Chaîne de Markov incluse

Soit (X_t) un processus markovien de saut, associé à $(T_n, Z_n), n \geq 0$. La suite $Z_n = X_{T_n}$ est une chaîne de Markov à temps discret. C'est une conséquence de la propriété de Markov forte de (X_t) . Elle est appelée **chaîne incluse** et vérifie que $Z_{n+1} \neq Z_n$, presque-sûrement, pour tout n . Sa matrice de transition se calcule aisément en fonction du générateur Q de X_t (grâce à (3.2.37)) :

$$\tilde{P}_{xy} = \begin{cases} (-Q_{xx})^{-1}Q_{xy}, & \text{si } y \neq x \\ 0, & \text{si } y = x. \end{cases}$$

Si on pose

$$S_n = q_{Z_{n-1}}(T_n - T_{n-1}),$$

où $q_x = -Q_{xx}$, et pour tout $t \geq 0$,

$$N_t = \sup\left\{n, \sum_{k=1}^n S_k \leq t\right\},$$

alors le processus (N_t) est un processus de Poisson d'intensité 1. En effet, il suffit d'appliquer le Corollaire 3.2.24 : la loi conditionnelle de $T_n - T_{n-1}$ sachant Z_{n-1} est une loi exponentielle de paramètre $q_{Z_{n-1}}$, et donc la loi de S_n est une loi exponentielle de paramètre 1. On utilise pour cela le fait que si U est une variable aléatoire exponentielle de paramètre λ , alors λU est une variable aléatoire exponentielle de paramètre 1.

Réciproquement, on peut définir un processus de Markov à temps continu à valeurs dans \mathbb{N} à partir de son générateur infinitésimal. Soit Q **une matrice de taux conservative** ou générateur infinitésimal, i.e. une matrice indexée par E telle que pour tous $x, y \in E$,

$$Q_{xy} \geq 0 \text{ si } y \neq x ; \quad -\infty < Q_{xx} = -\sum_{y \neq x} Q_{xy} \leq 0. \quad (3.2.42)$$

On pose alors $q_x = -Q_{xx}$ et on définit la matrice de transition P par $\tilde{P}_{xy} = \frac{Q_{xy}}{q_x}$ si $x \neq y$ et $\tilde{P}_{xy} = 0$ si $y = x$, avec la convention $\tilde{P}_{xy} = 0 \forall y$ si $Q_{xx} = 0$. A tout $x \in E$, on associe la chaîne $(Z_n)_n$ de matrice de transition \tilde{P} et on considère un processus de Poisson (N_t) d'intensité 1 indépendant de $(Z_n)_n$. On note $(T'_n)_n$ la suite des instants de saut de (N_t) et on définit pour $n \geq 1$

$$U_n = \frac{T'_n - T'_{n-1}}{q_{Z_{n-1}}} \quad ; \quad T_n = U_1 + \dots + U_n.$$

Si la condition de non-explosion (3.2.33) est satisfaite par $(T_n)_n$, alors

$$X_t = \sum_{n \geq 0} Z_n \mathbf{1}_{[T_n, T_{n+1}[}(t), \quad t \geq 0 \quad (3.2.43)$$

est un processus markovien de saut de générateur infinitésimal Q .

Construction algorithmique de $(X_t)_t$:

On peut toujours construire le processus $(X_t, t \geq 0)$ en itérant la procédure suivante.

- On démarre de $X_0 = x$, on attend un temps exponentiel U_1 de paramètre q_x .
- Etant donné $U_1 = t$, on fait sauter le processus à l'état y avec probabilité $\frac{Q_{xy}}{q_x}$.
- On réitère la procédure.

Les temps de vie exponentiels U_1, U_2, \dots sont appelés temps de séjour dans l'état Z_1, Z_2, \dots . Le processus minimal a une durée de vie finie si

$$T_\infty = \sum_{k \geq 1} U_k = \lim_n T_n < \infty.$$

Etudions maintenant la condition de non-explosion (3.2.33) pour $(T_n)_n$, qui assure que le processus est défini par (3.2.43) pour tout $t \in \mathbb{R}_+$, et prouve alors l'existence d'un processus markovien de saut de générateur infinitésimal Q .

Proposition 3.2.26 *La condition de non-explosion $\lim_n T_n = \infty$ p.s. est satisfaite si et seulement si*

$$\sum_{n \geq 0} q_{Z_n}^{-1} = +\infty \text{ p.s.} \quad (3.2.44)$$

Avant de prouver la proposition, énonçons tout de suite un corollaire immédiat.

Corollaire 3.2.27 *Pour qu'un générateur infinitésimal Q soit le générateur infinitésimal d'un processus markovien de saut vérifiant (3.2.42), il suffit que l'une des deux conditions suivantes soit satisfaite :*

- (i) $\sup_{x \in E} q_x < +\infty$.
- (ii) La chaîne de Markov $(Z_n)_n$ de matrice de transition \tilde{P} est récurrente.

Preuve. (de la proposition). Cela revient à montrer que si les (e_i) sont des variables aléatoires exponentielles indépendantes de paramètre q_i , alors presque-sûrement,

$$\mathcal{E} = \sum_i e_i = +\infty \Leftrightarrow \sum_i \frac{1}{q_i} = \infty.$$

En effet, $T_n - T_{n-1} = \frac{S_n}{q_{Z_{n-1}}}$, et les temps S_n sont indépendants. Ici on raisonne conditionnellement à la donnée des Z_n .

Si $\text{Card}\{i, q_i < 1\} = \infty$, alors $\sum_i e_i$ est supérieure à une somme infinie de variables aléatoires indépendantes de loi exponentielles de paramètre 1. Elle est donc infinie.

Si $\text{Card}\{i, q_i < 1\} < \infty$, introduisons la transformée de Laplace de \mathcal{E} . Elle vaut

$$\mathbb{E}(e^{-\lambda\mathcal{E}}) = \prod_i \mathbb{E}(e^{-\lambda e_i}) = \prod_i \frac{q_i}{\lambda + q_i}.$$

Elle sera nulle si et seulement si $\sum_i e_i = +\infty$ avec probabilité 1. Le produit ci-dessus est nul si et seulement si la somme $\sum_i \ln\left(1 + \frac{\lambda}{q_i}\right)$ est divergente. Or, la nature de la somme n'est pas modifiée si on enlève de la somme les termes d'indices i tels que $q_i < 1$. Puisque comme pour $q_i \geq 1$, on a $\frac{C^{te}(\lambda)}{q_i} \leq \ln\left(1 + \frac{\lambda}{q_i}\right) \leq \frac{\lambda}{q_i}$, on en déduit que $\sum_i \ln\left(1 + \frac{\lambda}{q_i}\right)$ converge si et seulement si $\sum_i \frac{1}{q_i} < \infty$. \square

3.3 Processus de branchement et de naissance et mort en temps continu

3.3.1 Processus de branchement en temps continu

Considérons une population composée d'un nombre dénombrable d'individus, où chaque individu donne naissance indépendamment des autres et au taux b à k , ($k \geq 1$) individus avec probabilité \tilde{p}_k , et meurt indépendamment des autres au taux constant d . Ainsi $\pi_k = b\tilde{p}_k$ est pour chaque individu le taux de production d'une couvée de taille k , et $a = b + d$ est le taux total de naissance et mort par individu.

Puisque chaque individu est remplacé par 0 avec probabilité $\frac{d}{a}$ et par $k + 1$ individus avec probabilité $\frac{\pi_k}{a}$, la chaîne incluse associée à ce schéma de naissance et mort est celle d'un processus de BGW avec fonction génératrice de reproduction

$$g(s) = \frac{1}{a} \left(d + \sum_{k \geq 1} \pi_k s^{k+1} \right), \quad s \in [0, 1].$$

Rappelons que la probabilité d'extinction s_0 de cette chaîne est la plus petite racine de l'équation

$$g(s) = s.$$

Le processus $(Z_t, t \geq 0)$ en temps continu décrit ci-dessus est un processus de Markov de saut pur. Les variables aléatoires décrivant les temps de reproduction et les temps de mort pour chaque individu suivent des lois exponentielles indépendantes. Nous savons que le minimum d'un nombre fini de variables exponentielles X_i indépendantes de paramètre a_i est une variable aléatoire exponentielle de paramètre la somme des a_i . (Voir Polycopié MAP 311). Ainsi, le processus de branchement en temps continu a des taux de transition (qui définissent le générateur (Q_{ij})) donnés par

$$\begin{cases} i \mapsto i + k & \text{au taux } i\pi_k, \quad k \geq 1, \\ i \mapsto i - 1 & \text{au taux } id. \end{cases} \quad (3.3.45)$$

Bien-sûr, $Q_{ij} = 0$, si $j < i - 1$. Rappelons qu'alors, comme vu au Théorème 3.2.22,

$$\begin{aligned} P_{i,i+k}(h) &= i\pi_k h + o(h), \quad \text{pour } k \geq 1, \\ P_{i,i-1}(h) &= id h + o(h), \\ P_{i,i}(h) &= 1 - i(b+d)h + o(h) = 1 - iah + o(h). \end{aligned}$$

Comme la chaîne de BGW incluse, le processus Z satisfait la propriété de branchement, à savoir que pour tout temps t , pour tout $i \geq 0$ et $s \in [0, 1]$,

$$\sum_{j=0}^{\infty} P_{i,j}(t) s^j = \left(\sum_{j=0}^{\infty} P_{1,j}(t) s^j \right)^i. \quad (3.3.46)$$

Définition 3.3.1 On appelle **processus de branchement continu** un tel processus.

Quand $\pi_1 = b$, c'est-à-dire si on ne peut avoir qu'un descendant à la fois, le processus est appelé **processus de branchement binaire** ou **processus de naissance et de mort linéaire**. Un tel processus modélise par exemple le mécanisme de division cellulaire.

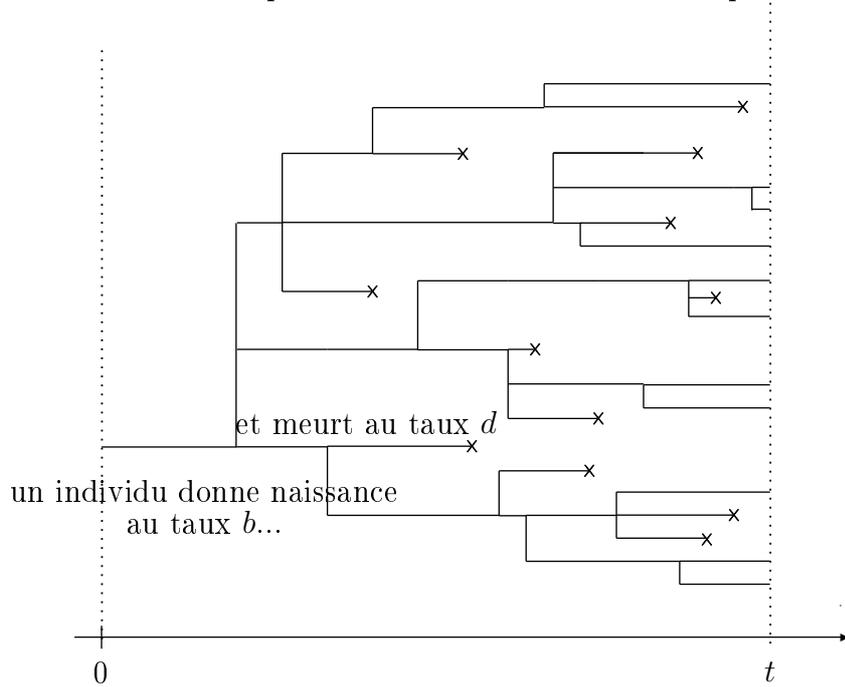
Si de plus, $d = 0$ (les individus ne meurent jamais), le processus est appelé **processus de fission binaire** ou **processus de Yule**.

Remarque 3.3.2 Donnons une autre interprétation du processus de branchement continu : chaque naissance d'une progéniture de k individus peut-être aussi considérée comme la naissance de $k + 1$ individus simultanément avec la mort de la mère. En d'autres mots, chaque individu vit durant un temps exponentiel de paramètre $a = b + d$ à la fin duquel il donne naissance à $k + 1$ individus avec probabilité

$$p_{k+1} = \frac{\pi_k}{a}, \quad k \geq 1; \quad p_1 = 0; \quad p_0 = \frac{d}{a}.$$

Signalons qu'il est possible de généraliser ce modèle en supposant que les temps de vie des individus suivent des distributions plus générales que des lois exponentielles. Le processus est alors appelé processus de Bellman-Harris. Nous n'étudierons pas ici ces processus.

FIG. 3.5 – Un processus de branchement en temps continu :



Nous avons

$$\begin{aligned} P_{i,j}(h) &= iap_{j-i+1}h + o(h), \quad \text{pour } j \geq i-1, j \neq i, \\ P_{i,i}(h) &= 1 - iah + o(h). \end{aligned}$$

En utilisant la section précédente, nous pouvons facilement écrire les équations de Kolmogorov pour le processus Z .

$$\frac{d}{dt}P_{i,j}(t) = -jaP_{i,j}(t) + a \sum_{1 \leq k \leq j+1, k \neq i} kp_{j-k+1}P_{i,k}(t), \quad (\text{progressive}) \quad (3.3.47)$$

$$\frac{d}{dt}P_{i,j}(t) = -iaP_{i,j}(t) + ia \sum_{k \geq i-1, k \neq i} p_{k-i+1}P_{k,j}(t) \quad (\text{rétrograde}), \quad (3.3.48)$$

avec les conditions initiales

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

En intégrant par une méthode de variation des constantes, nous pouvons écrire ces équations sous forme intégrale.

$$P_{i,j}(t) = \delta_{ij}e^{-jat} + a \sum_{k=1}^{j+1} \int_0^t e^{-ajs} k p_{j-k+1} P_{i,k}(t-s) ds, \quad (\text{progressive}) \quad (3.3.49)$$

$$P_{i,j}(t) = \delta_{ij}e^{-iat} + ia \sum_{k=i-1}^{\infty} \int_0^t e^{-ais} p_{k-i+1} P_{k,j}(t-s) ds. \quad (\text{rétrograde}) \quad (3.3.50)$$

Le processus peut exploser en un temps fini T_∞ avec probabilité positive, ou avoir un temps de vie infini. On a le théorème suivant.

Théorème 3.3.3 *Soit*

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

Alors le processus de branchement en temps continu a un temps de vie infini presque-sûrement si et seulement si $m = g'(1) < \infty$, ou si $m = \infty$ et

$$\int_{1-\varepsilon}^1 \frac{ds}{u(s)} = -\infty.$$

Preuve. Rappelons que s_0 est la plus petite racine de u . Nous excluons le cas trivial où $s_0 = 1$. (Dans ce cas, le processus BGW sous-jacent s'éteint, donc Z aussi).

Soit $h_i(t) = \mathbb{P}_i(T_\infty > t) = \sum_{j \geq 0} \mathbb{P}_i(Z_t = j)$. Remarquons que $h(t) := h_1(t)$ est décroissante en t , que $h(0^+) = 1$ et que $h(t) \in]s_0, 1]$. De plus, par la propriété de branchement, nous savons que $h_i(t) = h(t)^i$. En sommant les équations de Kolmogorov rétrogrades (3.3.49), nous obtenons que

$$h(t) = e^{-at} + a \int_0^t e^{-as} \sum_{k \geq 0} p_k h(t-s)^k ds, \quad t > 0.$$

En différenciant la dernière équation et en intégrant par parties, on obtient que

$$h'(t) = u(h(t)), \quad t > 0.$$

Supposons que le processus de branchement ait un temps de vie T_∞ fini avec probabilité positive de telle sorte que $h(t_0) < 1$ pour un temps t_0 . Fixons $\varepsilon \in]s_0, 1[$ et posons $F(x) = \int_\varepsilon^x \frac{ds}{u(s)}$. Comme la fonction $t - F(h(t))$ est de dérivée nulle, elle est constante, et $t - F(h(t)) = t_0 - F(h(t_0))$. En faisant tendre t vers 0, on obtient que $F(1^-)$ a une valeur finie. Alors, on en déduit que $\int_1^x \frac{ds}{u(s)}$ est fini, ce qui entraîne d'ailleurs que $m = \infty$. (On pourra faire un développement limité de $g(s)$ au voisinage de 1).

Réciproquement, supposons que $m = \infty$ et que $\int_1^x \frac{ds}{u(s)}$ converge. (On remarquera que si $m < \infty$, l'intégrale diverge forcément). On peut alors définir la fonction $G(x) = \int_1^x \frac{ds}{u(s)}$, pour $x \in]s_0, 1]$. En utilisant $h' = u(h)$, on obtient que $t - G(h(t)) = 0$, ce qui implique

$h(t) < 1$ dès que $t > 0$. □

Introduisons alors la probabilité d'extinction

$$q(t) = \mathbb{P}_1(T < t), \quad t \in]0, +\infty],$$

avec $q(0^+) = 0$ et $q(\infty) = s_0$.

Théorème 3.3.4 *La loi du temps d'extinction est donnée implicitement par*

$$\int_0^{q(t)} \frac{ds}{u(s)} = t, \quad t \geq 0.$$

Preuve. L'idée est la même que précédemment. Puisque $\mathbb{P}_i(Z_t = 0) = q(t)^i$, les équations de Kolmogorov rétrogrades conduisent pour $t > 0$ à

$$q(t) = a \int_0^t e^{-as} \left(\sum_{k \geq 0} p_{k+1} q(t-s)^{k+1} + d \right) ds.$$

Après dérivation et intégration par parties, nous obtenons que

$$q'(t) = u(q(t)), \quad t > 0. \tag{3.3.51}$$

Rappelons que $q(t) \in [0, s_0]$. Si nous posons $F(x) = \int_0^x \frac{ds}{u(s)}$ et si nous intégrons (3.3.51), nous obtenons finalement que $F(q(t)) = t$. □

3.3.2 Cas binaire

Dans le cas binaire, où $\pi_1 = b$, on a $u(s) = d - (b+d)s + bs^2$ et la probabilité d'extinction est $s_0 = \min(1, \frac{d}{b})$. Il est facile de vérifier que ce processus est critique si $b = d$ et est surcritique (resp. sous-critique) si $b > d$ (resp. $b < d$). Introduisons le taux de croissance $r = b - d$. On a alors

$$q(t) = \begin{cases} d(e^{rt} - 1)/(be^{rt} - d) & \text{si } b \neq d \\ bt/(1 + bt) & \text{si } b = d. \end{cases}$$

Dans le paragraphe sur les processus de naissance et mort, nous verrons de plus que le temps moyen d'extinction vaut

$$\mathbb{E}_1(T) = \frac{1}{b} \log \left(\frac{1}{1 - b/d} \right) \quad \text{si } b < d,$$

et est infini si $b \geq d$.

3.3.3 Extensions

Comme dans le cas du temps discret, nous allons pouvoir généraliser les modèles et considérer des processus de branchement avec *immigration* ou avec *croissance logistique*. Ces modèles sont de simples extensions du processus de branchement en temps continu, mais le premier ne va jamais s'éteindre et le deuxième ne va jamais exploser.

Immigration

Soit $\nu = (\nu_k)_{k \in \mathbb{N}}$ une mesure positive finie sur \mathbb{N} , ($\forall k, \nu_k \geq 0$ et $\sum_k \nu_k < \infty$), et soit $\rho = \sum_{k \geq 0} \nu_k$. Dans un modèle de branchement en temps continu avec immigration,

- Au taux ρ , des groupes d'immigrants arrivent dans la population.
- Un groupe est composé de k individus avec probabilité $\frac{\nu_k}{\rho}$.
- Tous les individus présents dans la population se reproduisent et meurent indépendamment suivant le schéma de branchement précédent.

Alors, le processus de branchement avec immigration a les taux de transition :

$$\begin{cases} i \rightarrow i + k & \text{au taux } i\pi_k + \nu_k \\ i \rightarrow i - 1 & \text{au taux } id. \end{cases}$$

Proposition 3.3.5 *Le processus de branchement avec immigration a un temps de vie infini presque-sûrement si et seulement si il en est de même du processus de branchement sous-jacent.*

Croissance logistique

Soit $c > 0$ appelé *intensité de la compétition*. Dans un modèle de branchement avec croissance logistique,

- Tous les individus présents dans la population se reproduisent et meurent indépendamment suivant le schéma de branchement précédent. (d est alors appelé taux de mort naturel).
- Un individu peut de plus subir la compétition avec les autres individus de la population, compétition pour les ressources par exemple. Le taux auquel il pourra être tué par un autre individu de la population, du fait de cette compétition, est égal à c .

Le processus de branchement avec croissance logistique, dit aussi processus de branchement logistique, a alors les taux de transition :

$$\begin{cases} i \rightarrow i + k & \text{au taux } i\pi_k \\ i \rightarrow i - 1 & \text{au taux } id + c i(i - 1). \end{cases}$$

On peut montrer le théorème suivant (Lambert 2005).

Théorème 3.3.6 *Si $d = 0$, le processus de branchement logistique Z_t converge en loi dans \mathbb{N}^* quand $t \rightarrow \infty$, alors que si $d \neq 0$, il s'éteint avec probabilité 1.*

Si $\sum_k \pi_k \ln(k) < \infty$, le processus de branchement logistique descend de l'infini, au sens où

$$\lim_{i \uparrow \infty} \mathbb{P}(Z_t = j | Z_0 = i) \text{ existe pour tout } j \geq 0, t > 0.$$

De plus, en désignant par T le temps d'extinction, on a que $\mathbb{E}_\infty(T) < \infty$.

3.3.4 Processus de naissance et mort

Définition 3.3.7 *Un processus de naissance et mort est un processus markovien de saut dont les amplitudes des sauts sont égales à ± 1 . Ses taux de transition sont donnés par*

$$\begin{cases} i \rightarrow i + 1 & \text{au taux } \lambda_i \\ i \rightarrow i - 1 & \text{au taux } \mu_i, \end{cases}$$

avec $\lambda_0 = \mu_0 = 0$.

C'est donc une généralisation d'un processus de branchement binaire pour des taux beaucoup plus généraux.

Exemples :

- 1) Le processus de Yule correspond à $\lambda_i = i\lambda$, $\mu_i = 0$.
- 2) Le processus de naissance et mort linéaire correspond à $\lambda_i = i\lambda$, $\mu_i = i\mu$.
- 3) Le processus de de naissance et mort avec immigration à $\lambda_i = i\lambda + \rho$, $\mu_i = i\mu$.
- 4) Le processus de de naissance et mort logistique à $\lambda_i = i\lambda$, $\mu_i = i\mu + c i(i - 1)$.

On peut alors montrer le théorèmes suivant. (Anderson, 1991)

Théorème 3.3.8 *Supposons que $\lambda_i > 0$ pour tout $i \geq 1$. Alors le processus de naissance et mort a un temps de vie infini presque-sûrement si et seulement si*

$$R := \sum_{i \geq 1} \left(\frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i \lambda_{i-1}} + \dots + \frac{\mu_i \cdots \mu_2}{\lambda_i \cdots \lambda_2 \lambda_1} \right) \text{ est infini.}$$

Remarque 3.3.9 Vous vérifierez que les 4 processus de naissance et mort mentionnés dans les exemples satisfont cette propriété et ont donc un temps de vie infini presque-sûrement.

3.3.5 Extinction

Posons

$$u_i := \mathbb{P}(\text{Extinction} | Z_0 = i),$$

qui est donc la probabilité que 0 soit atteint en temps fini. Appelons comme précédemment T le temps d'extinction et définissons également

$$\theta_i = \mathbb{E}(T \mathbf{1}_{\text{Extinction}} | Z_0 = i).$$

C'est le temps moyen pour atteindre l'extinction sur l'événement "Extinction", sachant que $Z_0 = i$. Alors nous avons les relations de récurrence suivantes, qui viennent du fait que les sauts sont d'amplitude 1. Pour tout $i \geq 1$,

$$\begin{cases} \lambda_i u_{i+1} - (\lambda_i + \mu_i) u_i + \mu_i u_{i-1} & = & 0 \\ \lambda_i \theta_{i+1} - (\lambda_i + \mu_i) \theta_i + \mu_i \theta_{i-1} & = & -u_i (\lambda_i + \mu_i). \end{cases}$$

Le seul cas non trivial est celui où tous les taux λ_i, μ_i sont non nuls ($i \geq 1$), ou quand ils sont non nuls jusqu'à un certain rang I . Définissons, pour un certain niveau I ,

$$u_i^{(I)} := \mathbb{P}_i(T < T_I).$$

Alors si on pose

$$U_I := \sum_{k=1}^{I-1} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k},$$

des calculs simples montrent que

$$u_i^{(I)} = (1 + U_I)^{-1} \sum_{k=i}^{I-1} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k},$$

pour $i \in \{1, \dots, I-1\}$.

En particulier, $u_1^{(I)} = \frac{U_I}{1+U_I}$.

Théorème 3.3.10 *Si $(U_I)_I$ tend vers l'infini quand $I \rightarrow \infty$, alors toutes les probabilités d'extinction sont égales à 1. Si $(U_I)_I$ converge vers une limite finie U_∞ , alors pour $i \geq 1$,*

$$u_i = (1 + U_\infty)^{-1} \sum_{k=i}^{\infty} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k}.$$

Application du Théorème 3.3.10 au processus de branchement (processus de naissance et mort) : Chaque individu naît à taux λ et meurt à taux μ . Nous avons donc un processus de branchement binaire où les individus vivent durant des variables exponentielles indépendantes de paramètre $\lambda + \mu$ et soit engendrent 2 individus avec probabilité $\frac{\lambda}{\lambda + \mu}$, soit meurent avec probabilité $\frac{\mu}{\lambda + \mu}$.

Alors, en appliquant les résultats précédents, on voit que quand $\lambda \leq \mu$, i.e. quand le processus est sous-critique ou critique, la suite $(U_I)_I$ tend vers l'infini quand $I \rightarrow \infty$ et on a extinction avec probabilité 1. Si en revanche $\lambda > \mu$, la suite $(U_I)_I$ converge vers $\frac{\mu}{\lambda - \mu}$, et un calcul simple montre que $u_i = (\lambda/\mu)^i$.

Application du Théorème 3.3.10 au processus de naissance et mort logistique
Supposons ici que les taux de croissance et de mort valent

$$\lambda_i = \lambda i ; \mu_i = \mu i + ci(i - 1). \quad (3.3.52)$$

Une explication biologique de ces hypothèses sera donnée dans la section suivante. Alors un calcul simple montre que dans ce cas, le processus de naissance et mort qualifié alors de logistique s'éteint presque-sûrement.

Revenons au cas général. On s'intéresse maintenant à l'espérance du temps d'extinction. Posons tout d'abord

$$\theta_i^{(I)} := \mathbb{E}_i(T, T < T_I), \quad i \leq I.$$

Alors par le théorème de convergence monotone, $\theta_i^{(I)}$ converge vers $\theta_i = \mathbb{E}_i(T \mathbf{1}_{Extinction})$ quand I tend vers ∞ . Soit

$$\rho_k := \frac{\lambda_1 \cdots \lambda_{k-1}}{\mu_1 \cdots \mu_k}.$$

Des calculs élémentaires conduisent alors au théorème suivant.

Théorème 3.3.11 *Sur l'événement $\{Extinction\}$, le temps d'extinction moyen est fini si et seulement si $\sum_k \rho_k u_k^2 < \infty$. Alors*

$$\theta_i = u_i \sum_{k=1}^{i-1} (1 + U_k) \rho_k u_k + (1 + U_i) \sum_{k \geq i} \rho_k u_k^2, \quad i \geq 1.$$

En particulier,

$$\mathbb{E}_1(T \mathbf{1}_{Extinction}) = \sum_{k \geq 1} \rho_k u_k^2.$$

On en déduit le corollaire suivant quand on a extinction presque-sûre (en posant $u_k = 1$).

Corollaire 3.3.12 *Quand $\mathbb{P}(Extinction) = 1$, le temps d'extinction moyen est fini si et seulement si $\sum_k \rho_k < \infty$. Alors*

$$\theta_i = \sum_{k=1}^{i-1} (1 + U_k) \rho_k + (1 + U_i) \sum_{k \geq i} \rho_k, \quad i \geq 1.$$

En particulier,

$$\mathbb{E}_1(T\mathbf{1}_{Extinction}) = \sum_{k \geq 1} \rho_k.$$

Application du théorème 3.3.11 au processus de branchement : Concentrons-nous sur le cas où on a extinction presque-sûre. Dans ce cas, $\lambda \leq \mu$, et $\rho_k = \frac{\lambda^{k-1}}{k\mu^k}$. On a alors

$$\mathbb{E}_1(T) = \lambda^{-1} \sum_{k \geq 1} k^{-1} \left(\frac{\lambda}{\mu}\right)^k = -\lambda^{-1} \ln\left(1 - \frac{\lambda}{\mu}\right), \quad \text{si } \lambda < \mu,$$

et $\mathbb{E}_1(T) = \infty$ si $\lambda = \mu$.

Preuve. (du Théorème 3.3.11). La relation de récurrence satisfaite par $\theta_i^{(I)}$ est

$$\lambda_i \theta_{i+1}^{(I)} - (\lambda_i + \mu_i) \theta_i^{(I)} + \mu_i \theta_{i-1}^{(I)} = -u_i^{(I)} (\lambda_i + \mu_i), \quad i \leq I.$$

Après quelques calculs, on obtient

$$\theta_i^{(I)} = (1 + U_i) \theta_1^{(I)} \sum_{k=1}^{i-1} \sigma_k^{(I)},$$

où

$$\sigma_k^{(I)} := \sum_{i=1}^k \frac{u_i^{(I)}}{\lambda_i} \prod_{j=i+1}^k \frac{\mu_j}{\lambda_j}.$$

(Un produit vide est égal à 1 par convention). On en déduit alors que

$$\sum_{k=1}^{i-1} \sigma_k^{(I)} = \sum_{k=1}^{i-1} \rho_k (U_i - U_k) u_k^{(I)}, \quad i \geq I.$$

De plus, comme $\theta_I^{(I)} = 0$, on obtient que

$$\theta_1^{(I)} = (1 + U_I)^{-1} \sum_{k=1}^{I-1} \rho_k (U_I - U_k) u_k^{(I)} = \sum_{k=1}^{I-1} \rho_k \left(u_k^{(I)}\right)^2.$$

Alors, en faisant tendre I vers l'infini, on obtient

$$\theta_1 = \sum_{k \geq 1} \rho_k u_k^2. \quad (3.3.53)$$

(Les deux termes peuvent être infinis). Si on fait de même pour $\theta_i^{(I)}$, on obtiendra de même

$$\theta_i = \mathbb{E}_i(T\mathbf{1}_{Extinction}) = (1 + U_i) \theta_1 - \sum_{k=1}^{i-1} \rho_k (U_i - U_k) u_k.$$

Si on remplace θ_1 par l'expression obtenue en (3.3.53), et en utilisant que quand $U_\infty < \infty$, $u_k = \frac{U_\infty - U_k}{1 + U_\infty}$, on obtient

$$\theta_i = u_i \sum_{k=1}^{i-1} \rho_k u_k (1 + U_k) + (1 + U_i) \sum_{k \geq i} \rho_k u_k^2.$$

Cette expression est aussi valable quand les probabilités d'extinction sont égales à 1. \square

3.4 Approximations continues : modèles déterministes et stochastiques

Comme dans le cas des dynamiques spatiales, nous pouvons observer que les calculs deviennent vite très compliqués pour les processus de naissance et mort que nous venons d'étudier. Nous allons maintenant introduire des approximations, soit déterministes, soit stochastiques, et retrouver ainsi des modèles classiques de Dynamique des populations. Ces différentes approximations vont donner des résultats qualitativement différents pour le comportement en temps long de la population et vont ainsi nous permettre de réfléchir à la pertinence du choix d'un modèle.

3.4.1 Approximations déterministes - Equations malthusienne et logistique

Supposons que le processus de naissance et mort soit paramétré par $N \in \mathbb{N}$, au sens où les taux de naissance λ_i^N et de mort μ_i^N dépendent du paramètre N . Ce paramètre N représente **la taille du système** qui suivant les cas, (et toutes ces interprétations sont liées entre elles de manière évidente), pourra décrire la taille de la population, l'extension du territoire ou la quantité de ressources disponibles.

Dans la suite de ce paragraphe, nous allons étudier le comportement asymptotique du processus de naissance et de mort quand $N \rightarrow \infty$. La taille de la population est de l'ordre de N . Ainsi pour obtenir une approximation raisonnable du processus, nous allons le renormaliser par le paramètre $\frac{1}{N}$ et nous allons étudier le processus

$$X_t^N = \frac{1}{N} Z_t^N \in \frac{1}{N} \mathbb{N},$$

où Z_t^N désigne le processus de naissance et mort avec les taux λ_i^N et μ_i^N . Les états pris par le processus X sont alors de la forme

$$\frac{i}{N}, \quad i \in \mathbb{N}.$$

Nous supposons tout d'abord que

$$\lim_{N \rightarrow \infty} X_0^N = \lim_{N \rightarrow \infty} \frac{1}{N} Z_0^N = x_0, \quad (3.4.54)$$

où la limite x_0 est supposée déterministe.

Une première remarque, fondamentale, est que les sauts du processus de naissance et mort ($Z_t^N, t \geq 0$) sont d'amplitude ± 1 . Ainsi, les sauts du processus ($X_t^N, t \geq 0$) sont d'amplitude $\pm \frac{1}{N}$, et vont donc tendre vers 0 quand $N \rightarrow \infty$. Si le processus ($X_t^N, t \geq 0$) converge vers un processus limite ($X_t, t \geq 0$), on peut assurer que *le processus* ($X_t, t \geq 0$) *sera continu en temps*. Remarquons également que les valeurs possibles prises par le processus X pourront être n'importe quel réel positif.

Le taux de croissance, quantité fondamentale pour comprendre le comportement en temps long de la population, est alors donné par

$$\lambda_i^N - \mu_i^N.$$

Nous allons tout d'abord supposer que

$$\begin{aligned} \text{Les taux } \frac{\lambda_i^N}{N} \text{ et } \frac{\mu_i^N}{N} \text{ sont uniformément bornés en } N, \\ \text{pour chaque } i \text{ tel que } \frac{i}{N} \text{ est uniformément borné en } N. \end{aligned} \quad (3.4.55)$$

Cette hypothèse sur les taux de naissance et de mort va, comme nous allons le voir par la suite, garantir que la limite du processus ($X_t^N, t \geq 0$) sera **une limite déterministe** (non aléatoire). De plus, nous allons supposer qu'il existe une fonction continue $H : \mathbb{R}_+ \rightarrow \mathbb{R}$ telle que

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} (\lambda_i^N - \mu_i^N) = H(x), \quad (3.4.56)$$

où $x \in \mathbb{R}_+$.

Deux cas particuliers sont particulièrement intéressants. Le cas de branchement linéaire : pour tout état x ,

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} (\lambda_i^N - \mu_i^N) = rx, \quad (3.4.57)$$

et le cas du branchement logistique : pour tout état x ,

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} (\lambda_i^N - \mu_i^N) = rx - cx^2, \quad (3.4.58)$$

où $r \in \mathbb{R}$ et $c > 0$.

Expliquons l'origine biologique de ces hypothèses. Le premier cas consiste à supposer que quand il y a un grand nombre d'individus, le taux de croissance est approximativement le même pour chaque individu. Le paramètre r représente alors le taux de croissance individuel dans cette population de taille infinie. Le taux de croissance d'une sous-population de i individus est alors approximativement proportionnel à i . L'exemple typique pour cette situation est celui du processus de branchement linéaire. Dans ce cas,

$$\lambda_i^N = \lambda i ; \quad \mu_i^N = \mu i,$$

où λ et μ sont des constantes strictement positives, et $r = \lambda - \mu$.

Ce modèle est obtenu dans un cas où l'on ne suppose *pas d'interaction* entre les individus, ce qui peut être considéré comme irréaliste. Le comportement limite est malthusien (voir Equation (3.4.60)), ce que l'observation de certaines populations stables contredit souvent. Sous une hypothèse de ressource globale fixée, il y a en général *compétition entre les individus* (pour le partage des ressources), ce qui va accroître le taux de mort dans les grandes populations et permettre ainsi de réguler, en cas de taux de croissance positif, l'évolution exponentielle de la population. Chaque individu est en compétition avec les $i - 1$ autres individus de la population. La biomasse de chaque individu, qui est l'énergie qu'il peut consacrer à la compétition, est proportionnelle à ses ressources individuelles, donc de la forme $\frac{c}{N}$. Ainsi, le taux de mort de la population est donné par

$$\mu_i^N = \mu i + \frac{c}{N} \times i(i - 1).$$

On va supposer que $\lambda_i^N = \lambda i$. Dans ce cas, le processus est appelé **processus de naissance et mort logistique**, nous l'avons déjà introduit en Section 3.3.5. Nous aurons

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} (\lambda_i^N - \mu_i^N) = rx - cx^2,$$

où $r = \lambda - \mu$. Le taux de croissance est ici quadratique en l'état de la population.

Remarquons que le choix d'un paramètre $c < 0$ peut aussi avoir un sens biologique : cela correspond à décrire des individus qui *coopèrent* : ils s'entraident pour subsister, et le terme quadratique est alors un terme qui accroît le taux de reproduction. Certains modèles plus sophistiqués décrivent des populations pour lesquels les individus coopèrent quand la taille de la population est inférieure à un certain seuil, et entrent en compétition quand la taille de la population dépasse ce seuil. Cet effet est appelé *effet Allee*. (cf. Kot [16]).

Equation limite

Supposons réalisées les hypothèses (3.4.55) et (3.4.56).

Etudions l'accroissement du processus $(X_t^N, t \geq 0)$ entre les temps t et $t+h$. En utilisant le Théorème 3.2.22, nous obtenons que pour tous $i \geq 1$ et k ,

$$\begin{aligned}\mathbb{E}(Z_{t+h}^N - Z_t^N | Z_t^N = i) &= ((i+1-i)\lambda_i^N + (i-1-i)\mu_i^N)h + o(h) \\ &= (\lambda_i^N - \mu_i^N)h + o(h).\end{aligned}$$

De plus,

$$\text{Var}(Z_{t+h}^N - Z_t^N | Z_t^N = i) = (\lambda_i^N + \mu_i^N)h - (\lambda_i^N - \mu_i^N)^2 h^2 + o(h).$$

Ainsi, nous en déduisons que

$$\begin{aligned}\mathbb{E}(X_{t+h}^N - X_t^N | Z_t^N = i) &= \frac{1}{N}(\lambda_i^N - \mu_i^N)h + \frac{1}{N}o(h), \\ \text{Var}(X_{t+h}^N - X_t^N | Z_t^N = i) &= \frac{1}{N^2}(\lambda_i^N + \mu_i^N)h + \frac{1}{N^2}o(h)\end{aligned}$$

On va supposer ici (on peut le montrer) que les infiniments petits $o(h)$ apparaissant dans les formules ci-dessus le sont uniformément en N . Bien-sûr, cela demanderait une preuve, mais qui dépasse le cadre du cours. Ce qu'il est important de comprendre ici est que si l'on accepte ce résultat, on obtient que quand N tend vers l'infini, et si le processus $(X_t^N, t \geq 0)$ converge vers un processus limite $(X_t, t \geq 0)$, le processus limite $(X_t, t \geq 0)$ est à trajectoires continues et vérifie : pour tout h

$$\begin{aligned}X_0 &= x_0, \\ \mathbb{E}(X_{t+h} - X_t | X_t) &= H(X_t)h + o(h), \\ \text{Var}(X_{t+h} - X_t | X_t) &= 0.\end{aligned}$$

Remarquons que les variances des accroissements sont nuls, et que la stochasticité disparaît donc quand N tend vers l'infini. En faisant tendre h vers 0, nous en déduisons alors que X est égal à son espérance, et est solution de l'équation différentielle

$$\frac{dX_t}{dt} = H(X_t) ; X_0 = x_0.$$

Nous justifions ainsi que dans ce cas la limite en "grande population" du processus Z^N initial se comporte comme la solution de l'équation différentielle

$$\dot{x}(t) = H(x(t)) ; x(0) = x_0. \quad (3.4.59)$$

L'étude des équilibres de cette équation est liée à la fonction H .

Equation malthusienne

Supposons que l'hypothèse (3.4.57) est satisfaite, à savoir que $H(x) = rx$. Alors, la limite en "grande population" est décrite par la solution de l'équation différentielle

$$\dot{x}(t) = r x(t) ; x(0) = x_0. \quad (3.4.60)$$

C'est l'**équation malthusienne**. Il est facile de décrire le comportement en temps long de la solution $x(t)$ en fonction du signe de r .

- $r > 0$, taux de croissance positif : $x(t) \rightarrow \infty$. La population explose.
- $r < 0$, taux de croissance négatif : $x(t) \rightarrow 0$. La population s'éteint.
- $r = 0$ taux de croissance nul. La population est stationnaire.

Equation logistique

Supposons maintenant que l'hypothèse (3.4.58) est satisfaite. Alors, la limite en "grande population" dans ce modèle limite avec compétition est décrite par la solution de l'équation différentielle

$$\dot{x}(t) = x(t)(r - cx(t)) ; x(0) = x_0, \quad (3.4.61)$$

appelée l'**équation logistique**.

Cette équation est très intéressante et célèbre en dynamique des populations dans le cas où le taux de croissance r est positif mais où la compétition va empêcher que la population n'explose.

Si $r > 0$, le terme quadratique (dû à la compétition) entraîne que la population se stabilise en temps long en une solution non triviale. En effet, il est facile de voir que l'équation a deux équilibres : 0 et $\frac{r}{c}$, et que 0 est instable.

Ainsi, la solution de l'équation logistique converge quand $t \rightarrow \infty$, dès que $x_0 \neq 0$, vers la quantité $\frac{r}{c} \neq 0$ appelée **capacité de charge**. Il est important de souligner la **différence de comportement en temps long** entre

- ce modèle limite (grande population) et le processus stochastique en petite population, puisque nous avons vu en Section 3.3.5 que le processus de naissance et de mort logistique s'éteint presque-sûrement.
- ce modèle logistique et le modèle malthusien.

La conclusion est que l'utilisation d'un tel modèle n'a de sens que sous certaines hypothèses de taille de population et ne prendra pas en compte les variations stochastiques dues aux petits effectifs pour certaines populations. Ainsi, en écologie, si l'on s'intéresse à l'extinction d'une population, il est clair que si la population passe en-dessous d'un certain effectif, le modèle déterministe perd son sens, et l'extinction devient une menace : elle devient probable même dans le cas logistique avec $r > 0$.

3.4.2 Approximation stochastique - Stochasticité démographique, Equation de Feller

Nous allons maintenant changer l'hypothèse (3.4.55) et supposer qu'il y a beaucoup plus de naissances et de morts, à savoir que les taux λ_i^N et μ_i^N sont maintenant de l'ordre de N . Plus précisément, nous supposerons l'hypothèse suivante : il existe $\gamma > 0$ tel que pour tout état x ,

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N^2} (\lambda_i^N + \mu_i^N) = \gamma x. \quad (3.4.62)$$

Cette hypothèse est en particulier satisfaite si $\lambda_i^N = \lambda i N$ et $\mu_i^N = \mu i N$, pour deux constantes positives λ et μ , et $\gamma = \lambda + \mu$. Les autres hypothèses restent inchangées, en particulier (3.4.56). Par rapport à l'étude précédente, le seul calcul qui change est le calcul des variances. Ici la limite des variances n'est pas nulle. Plus précisément, on obtient que

$$\text{Var}(X_{t+h} - X_t | X_t) = \gamma X_t h.$$

Rappelons que par ailleurs, $\mathbb{E}(X_{t+h} - X_t | X_t) = H(X_t)h + o(h)$. Nous supposons encore (3.4.54), en permettant d'avoir une limite X_0 qui peut être une variable aléatoire. On peut alors montrer (mais cela dépasse mathématiquement le cadre du cours) que la suite de processus $(X_t^N, t \geq 0)$ converge vers le processus $(X_t, t \geq 0)$ solution de l'équation différentielle stochastique

$$dX_t = H(X_t)dt + \sqrt{\gamma X_t}dB_t ; X_0. \quad (3.4.63)$$

En particulier sous les hypothèses $H(x) = rx$ ou $H(x) = rx - cx^2$, nous aurons

- Cas linéaire :

$$dX_t = r X_t dt + \sqrt{\gamma X_t} dB_t ; X_0. \quad (3.4.64)$$

- Cas logistique :

$$dX_t = (r - c X_t)X_t dt + \sqrt{\gamma X_t} dB_t ; X_0. \quad (3.4.65)$$

Dans les deux cas, $(B_t, t \geq 0)$ désigne un mouvement brownien. L'équation différentielle stochastique (3.4.64) est appelée **équation de Feller**, l'équation (3.4.65) est appelée **équation de Feller logistique**. Le cas général (3.4.63) a été introduit très récemment et porte le nom d'équation de Feller généralisée.

Nous avons ainsi, par notre étude, construit des solutions de ces équations par approximations. Nous pouvons également montrer que la solution de l'équation de Feller (3.4.64) est unique dans l'espace des processus de carré intégrable, en appliquant le Théorème 2.3.8. Pour l'équation de Feller logistique, c'est plus délicat car il y a un terme quadratique, et nous n'aborderons pas cette question ici, de même que pour l'équation de Feller généralisée.

L'étude de ces processus est difficile et fait plutôt l'objet d'un cours de master 2, mais l'important dans le cadre de notre cours est d'avoir compris le fait suivant. A partir du même modèle microscopique que précédemment, dans une hypothèse de grande population, mais sous l'hypothèse que les taux de naissance et de mort sont très grands, le "bruit" créé par les sauts permanents dus aux naissances et aux morts est tellement important qu'il va subsister à la limite. C'est ce qui explique l'apparition des termes stochastiques browniens et l'apparition d'une nouvelle **stochasticité démographique**.

Ainsi donc, nous avons introduit et justifié une grande gamme de modèles, probabilistes ou déterministes, continus en temps ou à saut, qui modélisent la dynamique d'une population.

3.5 Populations multi-type

Comme nous l'avions suggéré dans les exemples de la section 3.1.1, il est intéressant du point de vue pratique de considérer la dynamique d'une population composée de sous-populations d'individus de types différents. Le type d'un individu est défini comme un attribut (ou un ensemble d'attributs) qui reste fixé durant la vie de l'individu. Nous allons supposer que les individus ne peuvent prendre qu'un nombre fini de types. Des exemples classiques de types sont les classes de taille à la naissance, le sexe, le génotype de l'individu, son stade de maturité (juvénile, reproducteur) ... Le type peut affecter la distribution du nombre de descendants. Par exemple des individus juvéniles vont devenir des individus reproducteurs (aptés à la reproduction), ou rester juvéniles (inaptes à la reproduction), mais les individus reproducteurs produiront toujours des individus juvéniles. Les distributions des descendance seront donc différentes. (Elles n'ont pas le même support).

Un autre exemple de type est celui des génotypes : le génotype d'un parent affecte de manière évidente le génotype de ses descendants et peut aussi affecter leur effectif. Dans un modèle avec reproduction clonale (asexuée) et mutation, le parent de chaque génotype est supposé produire un nombre aléatoire de descendants avec une certaine répartition aléatoire des types. Avec une probabilité proche de 1, un descendant aura le même génotype que son parent mais avec une petite probabilité, il pourra être de génotype mutant. La sélection sur la fertilité ou les chances de survie peuvent aussi être modélisées par différentes formes de distributions des génotypes sur les descendants.

3.5.1 Le processus de branchement multi-type en temps discret

Supposons que les individus puissent prendre K types distincts. Le modèle le plus simple, qui généralise le modèle de BGW, consiste à décrire une population où pour chaque type $j \in \{1, \dots, K\}$, chaque individu de type j génère à la génération suivante, et indépendamment de tous les autres un K -uplet décrivant la répartition en les K types de ses descendants.

Exemple : La population est composée de deux types \heartsuit et \spadesuit . Nous avons 5 individus pour une certaine génération et à la génération suivante nous obtenons :

première génération : $\heartsuit \quad \heartsuit \quad \spadesuit \quad \spadesuit \quad \spadesuit$
 deuxième génération : $(\heartsuit, \heartsuit) \quad (\heartsuit, \spadesuit) \quad (\heartsuit, \spadesuit, \spadesuit) \quad (\heartsuit, \spadesuit, \spadesuit) \quad (\spadesuit)$.

Nous supposons de plus que pour chaque type j fixé, les k -uplets aléatoires issus de chaque individu de type j sont indépendants et de même loi, et que les descendance de tous les individus sont indépendantes entre elles.

Plus précisément, nous allons introduire le processus à valeurs vectorielles

$$X_n = (X_n^{(1)}, \dots, X_n^{(K)})^*,$$

(A^* désigne la matrice transposée de A), décrivant le nombre d'individus de chaque type à la génération n , qui est à valeurs dans \mathbb{N}^K . Supposons que la distribution initiale soit définie par le vecteur $(X_0^{(1)}, \dots, X_0^{(K)})^*$. La première génération sera alors définie par le vecteur $(X_1^{(1)}, \dots, X_1^{(K)})^*$, où pour $m \in \{1, \dots, K\}$,

$$X_1^{(m)} = \sum_{j=1}^K \sum_{i=1}^{X_0^{(j)}} \xi_j^{(m)}(i).$$

La variable aléatoire $\xi_j^{(m)}(i)$ représente le nombre d'enfants de type m issus du i -ième individu de type j . On a indépendance des descendance aléatoires correspondant à des types j différents. Pour j fixé, et conditionnellement à $X_0^{(j)}$, les vecteurs aléatoires $\xi_j^{(1)}, \dots, \xi_j^{(m)}, \dots, \xi_j^{(K)}$ sont indépendants. De plus, pour chaque j, m , chaque vecteur $(\xi_j^{(m)}(i), 1 \leq i \leq X_0^{(j)})$ est composé de variables aléatoires indépendantes et de même loi. La population à la génération 1 est par exemple donnée par

$$(X_1^{(1)}, \dots, X_1^{(K)})^* = \sum_{j=1}^K \sum_{i=1}^{X_0^{(j)}} (\xi_j^{(1)}(i), \dots, \xi_j^{(K)}(i))^*.$$

Dans l'exemple "coeur et pique", nous avons donc $X_0 = (2, 3)^*$ et $X_1 = (5, 6)^*$.

Remarque : Il faut bien comprendre que les descendants de chaque individu d'un certain type sont indépendants et de même loi, et que les descendance d'individus distincts à la n -ième génération sont indépendantes les unes des autres.

Définition 3.5.1 Ces processus de branchement sont appelés *processus de Bienaymé-Galton-Watson multi-type*.

Pour ces processus de Galton-Watson multi-type, les fonctions génératrices vont jouer le même rôle important que dans le cas monotype.

Pour tout $n \in \mathbb{N}$, tout type $j \in \{1, \dots, K\}$ et $\mathbf{s} = (s_1, \dots, s_K) \in [0, 1]^K$, nous allons définir la j -ième fonction génératrice $f_n^j(\mathbf{s})$ qui déterminera la distribution du nombre de descendants de chaque type (à la génération n) produits par une particule de type j .

$$f_n^j(\mathbf{s}) = \mathbb{E}(s_1^{X_n^{(1)}} \dots s_k^{X_n^{(K)}} \mid X_0^{(j)} = 1, X_0^{(m)} = 0, \forall m \neq j) = \sum_{i_1, \dots, i_K \geq 0} p_n^j(i_1, \dots, i_K) s_1^{i_1} \dots s_K^{i_K},$$

où $p_n^j(i_1, \dots, i_K)$ est la probabilité qu'un parent de type j produise i_1 descendants de type 1, ..., i_K descendants de type K à la génération n .

Par exemple, considérons une population de lynx. Ces animaux sont tout d'abord juvéniles. En grandissant les lynx deviennent pour la plupart aptes à la reproduction mais sont incapables de se reproduire s'ils ne se retrouvent dans une meute. Ils sont alors dits flottants. Dès lors qu'ils sont en meute, ils peuvent se reproduire. Toutefois, il arrive qu'un individu ne puisse s'intégrer à la meute et reste flottant. La population de lynx est donc composée de 3 types : juvénile (J), flottant (F), reproducteur (R). Un individu passe

- de l'état J à l'état F avec probabilité σ
 - de l'état J à l'état J avec probabilité $1 - \sigma$
 - de l'état F à l'état R avec probabilité σ'
 - de l'état F à l'état F avec probabilité $1 - \sigma'$
 - de l'état R à l'état F avec probabilité $1 - \rho$ (le lynx sort de la meute) et se reproduit avec probabilité ρ . Son nombre de petits suit alors une loi de Poisson de paramètre m .
- A la première génération, nous aurons donc

$$\begin{aligned} f_1^1(s_1, s_2, s_3) &= \sigma s_2 + (1 - \sigma) s_1 \\ f_1^2(s_1, s_2, s_3) &= \sigma' s_3 + (1 - \sigma') s_2 \\ f_1^3(s_1, s_2, s_3) &= e^{m(s_1-1)} \rho s_3 + (1 - \rho) s_2. \end{aligned}$$

Alors, avec un raisonnement analogue à celui du cas monotype, on peut montrer que

Proposition 3.5.2

$$\begin{aligned} f_n^j(\mathbf{s}) &= f_{n-1}^j(f_1^1(\mathbf{s}), f_1^2(\mathbf{s}), \dots, f_1^K(\mathbf{s})) \\ &= f_1^j(f_{n-1}^1(\mathbf{s}), f_{n-1}^2(\mathbf{s}), \dots, f_{n-1}^K(\mathbf{s})). \end{aligned}$$

Nous aurons extinction de la population si l'on a extinction de chaque type et donc si pour un temps n , l'on a $X_n^{(j)} = 0$, pour tous j . En raisonnant toujours comme dans le cas monotype, on peut montrer que

Théorème 3.5.3

$$\mathbb{P}(X_n^{(j)} = 0, \forall j \mid X_0^{(j)} = k_j, \forall j) = \prod_{j=1}^K (U_n^{(j)})^{k_j},$$

où $U_n^{(j)}, j \in \{1, \dots, K\}$ est l'unique solution du système

$$\begin{cases} U_0^{(j)} &= 0 \\ U_n^{(j)} &= f_1^j(U_{n-1}^{(1)}, \dots, U_{n-1}^{(K)}) \end{cases} \quad \forall j$$

Nous pourrions alors obtenir des résultats décrivant le comportement en temps long du processus, suivant le même raisonnement que dans le cas monotype, cf. Athreya-Ney [1] pour plus de détails. Toutefois, il nous faut comprendre dans ce cas multi-type ce qui correspond aux différents cas sous-critique, critique, surcritique développés dans le chapitre 3.1.2.

La reproduction moyenne sera décrite par une matrice

$$M = (m_{jk})_{j,k \in \{1, \dots, K\}}, \quad (3.5.66)$$

où

$$m_{kj} = \mathbb{E}(\xi_j^{(k)})$$

est le nombre moyen d'individus de type k issus d'un individu de type j . La somme des coefficients de la k -ième ligne représente donc le nombre moyen d'individus de type k et la somme des coefficients de la j -ième colonne représente le nombre moyen d'individus issus d'un individu de type j .

Proposition 3.5.4 *L'espérance du nombre d'individus de type k à la génération n vérifie*

$$\mathbb{E}(X_n^{(k)}) = \mathbb{E}(\mathbb{E}(X_n^{(k)} | X_{n-1})) = \sum_{j=1}^K m_{kj} \mathbb{E}(X_{n-1}^{(j)}), \quad (3.5.67)$$

et ainsi, on a l'égalité vectorielle

$$\mathbb{E}(X_n) = M^n \mathbb{E}(X_0). \quad (3.5.68)$$

On peut conclure qu'on aura sous-criticalité et extinction si la matrice $M^n \rightarrow 0$ et sur-criticalité si $M^n \rightarrow \infty$ (dans un sens à définir). La question qui se pose est alors de savoir si l'on peut caractériser la criticalité par un seul paramètre scalaire ρ .

De manière évidente, la matrice M est positive, au sens où tous ses termes sont positifs. Dans la suite, on fera l'hypothèse supplémentaire suivante (cf. Athreya-Ney [1], Haccou, Jagers, Vatutin [10]) :

Il existe un entier n_0 tel que tous les éléments de la matrice M^{n_0} sont strictement positifs (3.5.69)

De telles matrices ont des propriétés intéressantes (cf. Serre [20]).

Du point de vue biologique, cela veut dire que toute configuration initiale peut amener à toute autre composition en les différents types.

Définition 3.5.5 On dira que les vecteurs $u \in \mathbb{R}^K$ et $v \in \mathbb{R}^K$ sont des vecteurs propres à gauche et à droite de M associés à la valeur propre $\lambda \in \mathbb{C}$ si

$$u^*M = \lambda u^* \quad , \quad Mv = \lambda v,$$

ou encore

$$\lambda u_k = \sum_{j=1}^K u_j m_{jk} \quad , \quad \sum_{j=1}^K m_{kj} v_j = \lambda v_k.$$

Enonçons le théorème principal.

Théorème 3.5.6 (cf. Serre [20], [10]). *Théorème de Perron-Frobenius.*

Soit M une matrice carrée $K \times K$ à coefficients positifs et satisfaisant (3.5.69). Alors

1) il existe un unique nombre réel $\lambda_0 > 0$ tel que

- λ_0 est une valeur propre de M ,
- toute valeur propre λ de M (réelle ou complexe) est telle que

$$|\lambda| \leq \lambda_0.$$

La valeur propre λ_0 est appelée **valeur propre dominante** de M .

2) les coefficients qui composent les vecteurs propres à gauche u et à droite v , correspondant à la valeur propre λ_0 , peuvent être choisis tels que

$$\sum_{k=1}^K u_k = 1 \quad ; \quad \sum_{k=1}^K u_k v_k = 1.$$

Dans ces conditions, les vecteurs propres sont uniques. De plus,

$$M^n = \lambda_0^n A + B^n,$$

où $A = (v_k u_j)_{k,j \in \{1, \dots, K\}}$ et B sont des matrices telles que :

- $AB = BA = 0$,
- Il existe des constantes $\rho \in]0, \lambda_0[$ et $C > 0$ telles qu'aucun des éléments de la matrice B^n n'excède $C\rho^n$.

En appliquant ce théorème et si λ_0 désigne la valeur propre dominante de la matrice de reproduction moyenne, nous obtenons alors immédiatement la proposition suivante.

Proposition 3.5.7 *Le processus de branchement multi-type de matrice de reproduction M (satisfaisant (3.5.69)) est*

- *sous-critique* si $\lambda_0 < 1$,
- *critique* si $\lambda_0 = 1$,
- *surcritique* si $\lambda_0 > 1$.

On a

$$\mathbb{E}(X_n) \sim_{n \rightarrow \infty} \lambda_0^n A \mathbb{E}(X_0).$$

En particulier, si $\lambda_0 \leq 1$, le processus de branchement multi-type s'éteint presque-sûrement. Si la population initiale se réduit à un individu de type k , on a

$$\mathbb{E}(X_n^{(j)}) \sim_{n \rightarrow \infty} \lambda_0^n v_k u_j.$$

En temps long, le nombre moyen d'individus de type j divisé par la taille moyenne de la population sera donné par

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(X_n^{(j)})}{\mathbb{E}(|X_n|)} = \frac{u_j}{u_1 + \dots + u_K} = u_j,$$

où $|X_n|$ désigne le nombre total d'éléments de la population à la génération n . (On a supposé que $\sum_{k=1}^K u_k = 1$).

Ainsi, u_j représente la proportion d'individus de type j et v_k représente la fertilité d'un individu de type k . Du point de vue biologique, nous en déduisons que dans une population satisfaisant l'hypothèse (3.5.69), tous les types croissent (en moyenne) au même taux λ_0 . Les probabilités d'extinction peuvent en revanche différer, suivant le type de l'individu initial : plus fertile est le type de l'ancêtre, plus grande est sa chance de survie.

Exemple : Presque toutes les populations comportent des individus d'au moins deux stades de maturité : les individus juvéniles, inaptes à la reproduction et les individus reproducteurs. Les reproducteurs produisent des juvéniles et les juvéniles deviennent reproducteurs, ou restent à un état essentiellement juvéniles. Reprenons l'exemple de la population de lynx et assimilons les lynx flottants aux juvéniles. Nous avons alors une population formée de deux types 1 (juvéniles) et 2 (reproducteurs) avec une matrice de reproduction moyenne de la forme

$$M = \begin{pmatrix} s_1 & m \\ s_2 & 0 \end{pmatrix}.$$

On aura par exemple $\mathbb{E}(X_1^{(2)}) = m$. Il est facile de voir que la matrice M vérifie les hypothèses du théorème de Perron-Frobenius. Cherchons ici la valeur propre dominante. Soit $\lambda \in \mathbb{C}$.

$$\det(M - \lambda I) = -\lambda(s_1 - \lambda) - s_2 m.$$

Une analyse immédiate montre que

$$\lambda_0 > 1 \Leftrightarrow 1 - s_1 - m s_2 < 0,$$

$$\lambda_0 < 1 \Leftrightarrow 1 - s_1 - m s_2 > 0,$$

$$\lambda_0 = 1 \Leftrightarrow 1 - s_1 - m s_2 = 0.$$

Ainsi, la population de lynx va s'éteindre si $1 - s_1 - m s_2 \geq 0$, et se développer dans le cas contraire.

3.5.2 Les modèles proie-prédateur, systèmes de Lotka-Volterra

Comme dans le cas de populations monotypes, on peut développer des modèles de processus de branchement ou de naissance et de mort multi-type en temps continu. Le cas intéressant ici est le cas qui prend en compte les interactions entre les sous-populations des deux types.

Nous allons définir, pour une population initiale de taille N ,

- $r_i^{1,N}$: taux de croissance de la sous-population de type 1 dans l'état i ,
- $r_i^{2,N}$: taux de croissance de la sous-population de type 2 dans l'état i ,
- $\frac{c_{1,1}}{N} > 0$: taux de compétition entre deux individus de type 1,
- $\frac{c_{1,2}}{N} > 0$: taux de compétition d'un individu de type 2 sur un individu de type 1,
- $\frac{c_{2,1}}{N} > 0$: taux de compétition d'un individu de type 1 sur un individu de type 2,
- $\frac{c_{2,2}}{N} > 0$: taux de compétition entre deux individus de type 2.

Comme dans le cas monotype, nous supposons que pour tout état x ,

$$\lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} r_i^{1,N} = r_1 x ; \quad \lim_{N \rightarrow \infty, \frac{i}{N} \rightarrow x} \frac{1}{N} r_i^{2,N} = r_2 x , \quad (3.5.70)$$

où r_1 et r_2 sont deux nombres réels, ce qui veut dire qu'en l'absence de toute compétition, les deux populations auraient tendance à se développer à vitesse exponentielle.

Quand N tend vers l'infini et en adaptant les arguments de la section 3.4, nous pouvons montrer que le processus de naissance et mort $(\frac{1}{N} Z_t^N, t \geq 0) = ((\frac{1}{N} Z_t^{1,N}, \frac{1}{N} Z_t^{2,N}), t \geq 0)$, composé des deux sous-processus décrivant les tailles des sous-populations de type 1 et de type 2, (renormalisées par $\frac{1}{N}$), converge vers la solution déterministe $(x(t), t \geq 0) = ((x_1(t), x_2(t)), t \geq 0)$ du système suivant :

$$dx_1(t) = r_1 x_1(t) - c_{1,1} x_1(t)^2 - c_{1,2} x_1(t) x_2(t), \quad (3.5.71)$$

$$dx_2(t) = r_2 x_2(t) - c_{2,1} x_1(t) x_2(t) - c_{2,2} x_2(t)^2. \quad (3.5.72)$$

Ces systèmes ont été extrêmement étudiés par les biologistes théoriciens et par les spécialistes de systèmes dynamiques. Ils sont connus sous le nom de **Systèmes de Lotka-Volterra**. Bien-sûr, on peut généraliser l'hypothèse (3.5.70) comme en Section 3.4, et obtenir un système dynamique satisfait par $(x_t, t \geq 0)$ beaucoup plus compliqué.

Un cas particulier fondamental est celui des **modèles de proie-prédateur**.

Le modèle historique de prédation est dû à Volterra (1926) et de manière presque contemporaine à Lotka. Imaginons que les deux types des sous-populations sont respectivement une proie (un lapin) ou un prédateur de cette proie (un renard). Faisons les hypothèses suivantes :

- En l'absence de prédateurs, l'effectif de la population de proies croît exponentiellement,
- En l'absence de proies, l'effectif de la population de prédateurs décroît exponentiellement,
- Les taux de disparition des proies et de croissance des prédateurs sont proportionnels au nombre de rencontres entre une proie et un prédateur.

On a alors le modèle particulier suivant :

$$dx_1(t) = \alpha_1 x_1(t) - \beta_1 x_1(t)x_2(t), \quad (3.5.73)$$

$$dx_2(t) = -\alpha_2 x_2(t) + \beta_2 x_1(t)x_2(t), \quad (3.5.74)$$

où les paramètres α_1 , α_2 , β_1 et β_2 sont positifs. Pour l'étude de ces modèles, nous renvoyons en particulier au polycopié et au livre de Jacques Istas [12], [11]. Voir aussi Renshaw [19], Kot [16].

Chapitre 4

Génétique des populations

Ceux que nous appelions des brutes eurent leur revanche quand Darwin nous prouva qu'ils étaient nos cousins. George Bernard Shaw (1856 - 1950).

4.1 Quelques termes de vocabulaire

- Une cellule biologique est dite **haploïde** lorsque les chromosomes qu'elle contient sont chacun en un seul exemplaire. Le concept est généralement opposé à **diploïde**, terme désignant les cellules avec des chromosomes en double exemplaire. Chez les humains et la plupart des animaux, la reproduction génétique met en jeu une succession de phases diploïdes (phase dominante) et haploïde (phase de formation des gamètes).
- Un **gamète** est une cellule reproductrice de type haploïde qui a terminé la méiose.
- Un **locus**, en génétique, est un emplacement précis sur le chromosome. Il peut contenir un gène.
- Les **allèles** sont des versions différentes de l'information génétique codée sur un locus. Par exemple, les types "ridés" et "lisses" des pois, dans les expériences de Mendel, correspondent à des allèles distincts. Dans le cas du gène déterminant le groupe sanguin, situé sur le chromosome 9 humain, l'un des allèles code le groupe A, un autre pour le groupe B, et un troisième allèle détermine le groupe O.
- L'**avantage sélectif**, ou **fitness** d'un allèle est une mesure qui caractérise son aptitude à se transmettre, et qui dépend ainsi de l'aptitude qu'il confère à son porteur à se reproduire, survivre...

4.2 Le cycle de la reproduction

4.3 Introduction

Nous nous intéressons ici à la reproduction des individus et principalement à la transmission de leurs allèles.

Dans une population haploïde, la proportion des individus portant un allèle donné fournit directement la fréquence de cet allèle dans la population. Mais dans les populations diploïdes, les gènes sont associés par paires dans les individus. Il est alors nécessaire de distinguer 2 types de fréquences pour décrire la composition génétique de la population à un locus considéré : les fréquences génotypiques qui sont les fréquences des différents génotypes à un locus considéré, et les fréquences alléliques qui sont les fréquences des différents allèles au locus considéré. En général, les fréquences génotypiques ne peuvent pas se déduire des fréquences alléliques sauf dans les hypothèses simplificatrices du modèle idéal de Hardy-Weinberg que nous allons développer ci-dessous. Supposons par exemple que l'on a deux allèles A et a . Une population avec 50% de gènes A et 50% de gènes a peut-être constituée uniquement d'homozygotes AA et aa , ou uniquement d'hétérozygotes Aa , ou de diverses proportions entre ces 3 génotypes.

4.3.1 Un modèle idéal de population infinie : le modèle de Hardy-Weinberg

Le premier modèle publié concernant la structure génotypique d'une population l'a été simultanément par Hardy et par Weinberg en 1908. Dans ce modèle, on fait un certain nombre d'hypothèses permettant de simplifier les calculs, et qui le rendent essentiellement déterministe (nombre infini de gamètes, population de taille infinie) :

- Les gamètes s'associent au hasard indépendamment des gènes considérés (hypothèse de panmixie). Cette hypothèse revient à dire que l'on considère une urne avec une infinité de gamètes qui sont appariés au hasard, sans tenir compte du sexe de l'individu.
- La population a une taille infinie. Par la loi des grands nombres, on remplace la fréquence de chaque allèle par sa probabilité.
- La fréquence des gènes n'est pas modifiée d'une génération à la suivante par mutation, sélection, migration.

Sous ces hypothèses, supposons qu'en un locus, les probabilités des allèles A et a soient p et $q = 1 - p$. Alors, à la deuxième génération, après appariement d'une gamète mâle et d'un gamète femelle, on a le génotype AA avec probabilité p^2 , le génotype aa avec probabilité q^2 et le génotype Aa avec probabilité $2pq$. Cette structure génotypique est connue sous le nom de structure de Hardy-Weinberg. Mais alors, puisque chaque individu a deux copies de chaque gène, la probabilité d'apparition de l'allèle A dans la population, à la deuxième génération sera $\frac{2p^2+2pq}{2} = p^2 + pq = p$. De même, la fréquence de l'allèle a sera q .

Ainsi, étant donné les hypothèses faites, on peut énoncer la loi de Hardy-Weinberg : **Dans une population isolée d'effectif illimité, non soumise à la sélection et dans laquelle il n'y a pas de mutation, les fréquences alléliques restent constantes.** Si les accouplements sont panmictiques, les fréquences génotypiques se déduisent directement des fréquences alléliques et restent donc constantes, et égales à $p^2, 2pq, q^2$.

Il suffit donc de connaître à chaque génération les probabilités de réalisation de l'allèle A et de celle de l'allèle a , et du fait des hypothèses de population d'individus infinie et de nombre de gamètes infini, le modèle ne présente pas de variation aléatoire. Il se réduit à un calcul de probabilité sur les fréquences alléliques, et l'on peut se ramener à un modèle haploïde.

4.4 Population finie : le modèle de Wright-Fisher

4.4.1 Modèle de Wright-Fisher

Alors que dans la population de taille infinie les fréquences alléliques sont stables au cours des générations en l'absence de sélection et de mutation (loi des grands nombres), les fréquences alléliques varient aléatoirement dans des populations de taille finie. (Cela est dû à la variabilité dans la distribution des gènes d'une génération à l'autre). Pour permettre un traitement mathématique pas trop compliqué, le modèle de Wright-Fisher modélise la transmission des gènes d'une génération à l'autre de manière très schématique. Il fait l'hypothèse de générations séparées, ce qui est une simplification considérable du cycle de reproduction. Ici, une population de M individus est représentée par un vecteur de $N = 2M$ allèles.

On notera par $n \in \mathbb{N}$ les indices de générations. Dans le modèle de Wright-Fisher, le parent de chaque individu de la génération $n + 1$ est distribué uniformément dans la n -ième génération. On suppose également que la population d'individus est de taille finie et constante M (On cultive des petits pois et à chaque génération on en garde M). Chaque individu est caractérisé par deux types (ou allèles) A et a qu'il transmet par hérédité. On a donc $N = 2M$ allèles. Le modèle de Wright-Fisher est un **modèle neutre**. Cela veut dire qu'il n'y a pas d'avantage sélectif associé à l'un des deux types (qui favoriserait la reproduction d'un des allèles). Il n'y a pas de mutation (modèle simpliste dans la pratique mais très intéressant dans une première approximation). Nous voulons étudier la fréquence allélique des deux allèles A et a à un locus donné, au fil des générations. Supposons que l'on connaisse les fréquences alléliques à une certaine génération n . Quelles vont-elles être à la génération $n + 1$ suivante ? On suppose que les individus peuvent se reproduire de manière indépendante les uns des autres, et que la nouvelle génération est formée de N gènes choisis uniformément dans une urne gamétique que l'on suppose infinie (il existe des milliers de gamètes), et dans laquelle la répartition allélique est celle de la génération n . Ainsi, le modèle qui consiste à choisir M individus d'allèles A ou a à la génération $n + 1$ (et donc $N = 2M$ allèles), et à compter ceux d'allèles A , devrait suivre

une loi hypergéométrique. Mais l'hypothèse d'urne infinie de gamètes permet d'approcher cette loi hypergéométrique (tirage simultané) par une loi binomiale qui ne prend plus en compte le nombre d'individus mais seulement les proportions des 2 allèles dans l'urne et qui correspond à un tirage avec remise. (Voir l'étude des modèles d'urnes, cours MAP 311 "Aléatoire", S. Méléard, Section 2.2.3).

Du fait de la neutralité du modèle, tous les individus sont **échangeables**, et seule la répartition des allèles a une importance. Définissons

$$X_n^N = \text{nombre d'individus de type } A \text{ à la génération } n.$$

On a le résultat suivant.

Proposition 4.4.1 *Pour tout $n \in \mathbb{N}$, pour tous i et j dans $\{0, \dots, N\}$,*

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}. \quad (4.4.1)$$

Preuve. Par définition, pour tout $n \in \mathbb{N}$, la variable X_n^N prend ses valeurs dans $\{0, \dots, N\}$. Il est clair que $\mathbb{P}(X_{n+1}^N = 0 | X_n^N = 0) = 1$, par définition, et de même, $\mathbb{P}(X_{n+1}^N = N | X_n^N = N) = 1$. Plus généralement, soit $i \in \{1, \dots, N-1\}$. Conditionnellement au fait que $X_n^N = i$, la fréquence de l'allèle A à la génération n est $\frac{i}{N}$ et celle de l'allèle a est $\frac{N-i}{N} = 1 - \frac{i}{N}$. Ainsi, à la génération suivante, chaque individu tire son parent au hasard, sans se préoccuper des autres individus. Cela correspond donc à un tirage avec remise, et le nombre d'individus d'allèle A à la $(n+1)$ -ième génération, sachant que $X_n^N = i$, suit alors une loi binomiale $\mathcal{B}(N, \frac{i}{N})$. Ainsi,

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}.$$

□

Remarque 4.4.2 Dans les modèles non neutres que nous verrons dans la suite (avec mutation, ou sélection), les individus ne seront plus échangeables, et la probabilité de "tirer" un individu d'allèle A connaissant $X_n^N = i$ ne sera plus uniforme.

Remarque 4.4.3 *D'après la proposition 4.4.1, si on note $p_i = \frac{i}{N}$, on aura que*

$$\mathbb{E}(X_{n+1}^N | X_n^N = i) = Np_i = i \quad \text{et donc} \quad \mathbb{E}(X_{n+1}^N - X_n^N | X_n^N = i) = 0 \quad (4.4.2)$$

$$\text{Var}(X_{n+1}^N - X_n^N | X_n^N = i) = Np_i(1 - p_i) = i\left(1 - \frac{i}{N}\right). \quad (4.4.3)$$

Nous nous intéressons à la transmission de l'allèle A . Les questions importantes sont les suivantes : l'allèle A va-t-il envahir toute la population ? Ou l'allèle a ? En combien de temps ? Aura-t-on co-existence ?

Théorème 4.4.4 *La suite $(X_n^N)_{n \in \mathbb{N}}$ est une chaîne de Markov à espace d'états fini et est une martingale bornée. La matrice de transition P de la chaîne est donnée pour tous i et j dans $\{0, \dots, N\}$ par*

$$P_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}. \quad (4.4.4)$$

Les états 0 et N sont deux états absorbants pour la chaîne.

Preuve. 1) Par définition, le processus $(X_n^N, n \geq 0)$ prend ses valeurs dans l'ensemble fini $\{0, \dots, N\}$. Il est évident que ce processus définit une chaîne de Markov puisque la loi de X_{n+1}^N conditionnellement à la connaissance de X_n^N est parfaitement définie par le calcul précédent. La forme de la matrice de transition a été calculée à la proposition 4.4.1.

Par ailleurs, si \mathcal{F}_n désigne la filtration engendrée par X_0^N, \dots, X_n^N , on a

$$\begin{aligned} \mathbb{E}(X_{n+1}^N | \mathcal{F}_n) &= \sum_{j=0}^N j \mathbb{P}(X_{n+1}^N = j | \mathcal{F}_n) \\ &= \sum_{j=0}^N j \mathbb{P}(X_{n+1}^N = j | X_n^N) = \sum_{j=0}^N j \binom{N}{j} \left(\frac{X_n^N}{N}\right)^j \left(1 - \frac{X_n^N}{N}\right)^{N-j} \\ &= N \frac{X_n^N}{N} = X_n^N. \end{aligned}$$

Ainsi, le processus $(X_n^N)_n$ est une martingale, bornée par N .

2) Il est évident que les états 0 et N sont absorbants. (voir (4.4.1)). □

Remarque 4.4.5 Remarquons que le processus $(X_n^N)_n$ est une martingale, et donc la taille de la population d'allèle A est constante en moyenne si et seulement si le modèle est neutre : il n'y a pas d'avantage sélectif pour un allèle. On verra en Section 4.4.3 comment la sélection influe sur cette dynamique.

Nous allons étudier comment les allèles se fixent dans la population d'individus de taille fixée N , en temps long. Nous allons donc faire tendre n vers l'infini.

Théorème 4.4.6 *Quand le nombre de générations n tend vers l'infini, la suite de variables aléatoires $(X_n^N)_n$ converge presque-sûrement vers une variable aléatoire X_∞^N et*

$$X_\infty^N \in \{0, N\}. \quad (4.4.5)$$

De plus,

$$\mathbb{P}_i(X_\infty^N = N) = \frac{i}{N} = \frac{1}{N} \mathbb{E}(X_\infty^N | X_0^N = i). \quad (4.4.6)$$

Remarquons que $\mathbb{P}_i(X_\infty^N = N)$ est la probabilité de fixation de l'allèle A dans la population et que $\mathbb{P}_i(X_\infty^N = 0)$ est sa probabilité de disparition et de fixation de l'allèle a .

Preuve. Puisque $(X_n^N)_n$ est une martingale bornée, elle converge presque-sûrement quand $n \rightarrow \infty$ vers une variable aléatoire X_∞^N . Le processus $(X_n^N)_n$ est une chaîne de Markov d'espace d'état fini. On sait qu'alors les deux points absorbants sont des états récurrents positifs. Tous les autres états sont transients. En effet, si $i \in \{1, \dots, N-1\}$,

$$p_{i0} = \mathbb{P}(X_1^N = 0 | X_0^N = i) = \left(1 - \frac{i}{N}\right)^N > 0$$

d'après (4.4.1), donc i mène à 0 mais 0 ne mène pas à i puisque 0 est absorbant. On a donc deux classes de récurrence $\{0\}$ et $\{N\}$ et une classe transiente $\{1, \dots, N-1\}$ dont la chaîne sort presque-sûrement à partir d'un certain rang. Ainsi, X_∞^N prend ses valeurs dans $\{0, N\}$. Introduisons le temps de fixation

$$\tau = \inf\{n \geq 0; X_n^N = 0 \text{ ou } X_n^N = N\} = T_0 \wedge T_N, \quad (4.4.7)$$

si $T_m = \inf\{n, X_n^N = m\}$. le temps d'arrêt τ est donc fini presque-sûrement. Notons \mathbb{P}_i la loi de la chaîne issue de i . On a alors

$$\mathbb{E}_i(X_\tau^N) = \mathbb{E}_i(X_0^N) = i.$$

En effet, par la propriété de martingale, on sait que pour tout n , $\mathbb{E}_i(X_n^N) = \mathbb{E}_i(X_0^N) = i$. Utilisant le fait que $X_n^N = X_\tau^N$ pour $n \geq \tau$, on a

$$i = \mathbb{E}_i(X_n^N) = \mathbb{E}_i(X_\tau^N \mathbf{1}_{\tau \leq n}) + \mathbb{E}_i(X_n^N \mathbf{1}_{\tau > n}).$$

Quand n tend vers l'infini, et puisque $|X_n| \leq N$, et $\tau < \infty$ p.s., on conclut par convergence dominée que le premier terme du membre de droite converge vers $\mathbb{E}_i(X_\tau^N)$ et le deuxième vers 0. On obtient donc le résultat, puisque cela entraîne en particulier que

$$P_i(X_\infty^N = N) = \frac{i}{N}.$$

□

Pour avoir une idée du temps que la fixation met pour avoir lieu, étudions la probabilité que deux copies d'un même locus choisies au hasard (sans remise) portent des allèles différents (**hétérozygotie**). Soit $h(n)$ cette probabilité, dans le cas où ces copies sont choisies dans la n -ième génération. Calculons tout d'abord la probabilité d'avoir l'hétérozygotie, conditionnellement à la connaissance de $X_n^N = i$. On a alors

$$H_n = \frac{\binom{i}{1} \binom{N-i}{1}}{\binom{N}{2}} = \frac{2i(N-i)}{N(N-1)}.$$

Plus généralement, H_n est la variable aléatoire

$$H_n = \frac{\binom{X_n^N}{1} \binom{N-X_n^N}{1}}{\binom{N}{2}} = \frac{2X_n^N(N-X_n^N)}{N(N-1)}.$$

Alors

Proposition 4.4.7

$$h(n) = \mathbb{E}(H_n) = \left(1 - \frac{1}{N}\right)^n \mathbb{E}(H_0). \quad (4.4.8)$$

Preuve. Il est pratique de numéroter les N copies d'un locus et d'en parler comme d'individus. Nous sommes alors ramenés à l'étude de la généalogie de ces individus. (Nous allons remonter le temps). En effet, les deux individus au temps n seront hétérozygotes s'ils n'ont pas eu d'ancêtre commun, et si au temps 0, leurs parents étaient également hétérozygotes.

Supposons qu'au temps n , on choisisse deux individus numérotés $x_1(0)$ et $x_2(0)$ distincts. Les individus $i = 1, 2$ sont chacun descendants d'individus de marques $x_i(1)$ au temps $n-1$, qui sont descendants d'individus de marques $x_i(2)$ au temps $n-2$ Ainsi la suite $x_i(m), 0 \leq m \leq n$ décrit la généalogie de $x_i(0)$, c'est-à-dire la suite de ses ancêtres si on a changé le sens du temps. Voir Figure 4.1.

Remarquons que si $x_1(m) = x_2(m)$, les deux individus ont eu le même ancêtre m générations plus tôt, et alors $x_1(l) = x_2(l)$ pour $m \leq l \leq n$. Si $x_1(m) \neq x_2(m)$, alors puisque les choix des deux parents des individus sont fait indépendamment l'un de l'autre, la probabilité qu'ils aient le même ancêtre commun à la $(m+1)$ -ième génération est $\frac{1}{N}$ (probabilité que $x_1(m)$ et $x_2(m)$ aient le même parent), et on a que $x_1(m+1) \neq x_2(m+1)$ avec probabilité $1 - \frac{1}{N}$. Ainsi, par récurrence, on montre que la probabilité d'avoir $x_1(n) \neq x_2(n)$, c'est à dire que tous les ancêtres des deux individus $x_1(0)$ et $x_2(0)$ sont différents jusqu'à la n -ième génération (passée) est égale à $\left(1 - \frac{1}{N}\right)^n$. Dans ce cas, les deux lignées sont disjointes jusqu'à cette génération passée n . Conditionnellement à cet événement, les deux individus $x_1(n)$ et $x_2(n)$ sont donc choisis au hasard dans la population au temps réel 0 et la probabilité qu'ils soient différents est la fréquence $H_0 = h(0)$. On conclut ainsi la preuve. \square

Remarquons que si l'on part avec une seule copie d'allèle A , alors $X_0^N = 1$, et $H_0 = \frac{2}{N}$. Ainsi, $h(0) = \frac{2}{N}$, et dans ce cas, grâce à (4.4.8), la probabilité qu'à la génération n , deux copies choisies au hasard aient deux allèles différents vaut $\frac{2}{N} \left(1 - \frac{1}{N}\right)^n$. Cela nous donne la vitesse de décroissance de l'hétérozygotie en fonction du temps.

4.4.2 Modèle de Wright-Fisher avec mutation

Supposons maintenant qu'au cours de la reproduction, on ait des mutations de l'allèle A vers l'allèle a avec probabilité α_1 et de l'allèle a vers l'allèle A avec probabilité α_2 . On

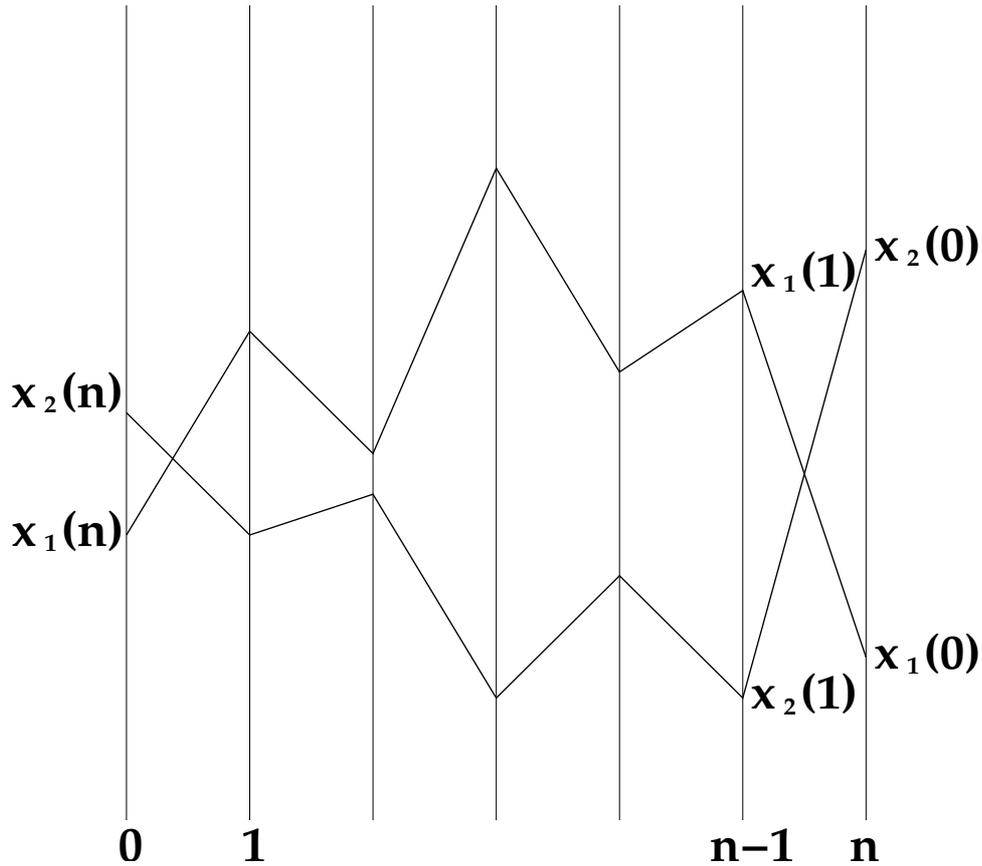


FIG. 4.1 –

suppose toujours que le modèle est neutre, c'est à dire que les deux allèles ont le même avantage sélectif.

Alors si l'on désigne toujours par $(X_n^N)_n$ le processus qui décrit le nombre d'allèles A à la génération n , on sait que la loi de X_{n+1}^N est une loi binomiale de paramètres qui dépendent de la répartition de l'allèle A à la génération n . Si on a i copies de l'allèle A à la génération n , cette répartition sera donnée par

$$p_i = \frac{i(1 - \alpha_1) + (N - i)\alpha_2}{N}.$$

Alors X_{n+1}^N suit une loi binomiale $\mathcal{B}(N, p_i)$, et on a

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} (p_i)^j (1 - p_i)^{N-j}, \quad (4.4.9)$$

avec $p_i = \frac{i(1 - \alpha_1) + (N - i)\alpha_2}{N}$.

4.4.3 Modèle de Wright-Fisher avec sélection

Nous supposons ici que l'allèle A a un avantage sélectif sur l'allèle a . Cela veut dire qu'une copie portant l'allèle A a plus de chance de se répliquer. Soit $s > 0$ le paramètre décrivant cet avantage. Si l'on suppose qu'à la génération n il y a i individus possédant l'allèle A , l'avantage sélectif sera modélisé par le fait que la probabilité d'obtention de l'allèle A sera donnée par

$$p_i = \frac{i(1+s)}{i(1+s) + N - i}. \quad (4.4.10)$$

On aura alors

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \binom{N}{j} (p_i)^j (1 - p_i)^{N-j}. \quad (4.4.11)$$

Plus généralement, on peut définir un modèle de Wright-Fisher avec sélection et mutation, où

$$p_i = \frac{(1+s)(i(1-\alpha_1) + (N-i)\alpha_2)}{(1+s)(i(1-\alpha_1) + (N-i)\alpha_2) + i\alpha_1 + (N-i)(1-\alpha_2)}. \quad (4.4.12)$$

Remarquons que la différence $p_i - \frac{i}{N}$ de la fraction d'allèles A à la fin du cycle de la n -ième génération exprime la différence due à la sélection et à la mutation dans le modèle déterministe de population infinie. Les fluctuations statistiques dues à la taille finie de la population apparaissent à travers le comportement aléatoire de la chaîne de Markov, qui se traduit par les probabilités de transitions P_{ij} . Il est rarement possible, pour une telle chaîne, de calculer les probabilités d'intérêt. Nous allons voir dans le paragraphe suivant que si la taille N de la population est grande, et si les effets individuels des mutations et de la sélection sont faibles, alors le processus peut être approché par un processus de diffusion pour lequel il sera plus facile d'obtenir des résultats quantitatifs.

4.5 Modèles démographiques de diffusion

4.5.1 Diffusion de Fisher-Wright

Nous considérons donc ici une approximation du modèle de Wright-Fisher, dans le cas où la taille de la population N tend vers l'infini. On suppose que $\frac{1}{N}X_0^N$ a une limite z quand $N \rightarrow \infty$. Dans ce cas, le nombre d'individus d'allèle A est de l'ordre de la taille de la population : on peut voir que l'espérance et la variance de la chaîne calculées à la remarque 4.4.3 sont de l'ordre de N . Pour pouvoir espérer une limite intéressante, nous allons considérer le processus $(\frac{1}{N}X_n^N, n \geq 0)$. Alors, si $Y_n^N = \frac{1}{N}X_n^N$, on aura

$$\begin{aligned} \mathbb{E}(Y_{n+1}^N | Y_n^N = x) &= \frac{1}{N} \mathbb{E}(X_{n+1}^N | X_n^N = Nx) = x \quad \text{et donc} \quad \mathbb{E}(Y_{n+1}^N - Y_n^N | Y_n^N = x) = 0 \\ \text{Var}(Y_{n+1}^N - Y_n^N | Y_n^N = x) &= \frac{1}{N^2} \text{Var}(X_{n+1}^N - X_n^N | X_n^N = Nx) = \frac{1}{N^2} Nx(1-x). \end{aligned}$$

Nous voyons donc que si $N \rightarrow \infty$, l'espérance et la variance des accroissements étant asymptotiquement nuls, le processus semble rester constant. L'échelle de temps n'est donc pas la bonne pour observer quelque chose, les événements de reproduction n'étant pas dans la même échelle que la taille de la population (fluctuations beaucoup trop petites dans cette échelle de taille). On va alors accélérer le temps, c'est à dire considérer une unité de temps qui dépend de la taille de la population. On peut soit dire qu'on regarde le processus dans une échelle de temps très longue, soit qu'on accélère le processus de reproduction. On va donc poser $n = [Nt]$, pour $t \in [0, T]$, $[x]$ désignant la partie entière de x . Reprenons les calculs précédents en posant

$$Z_t^N = \frac{1}{N} X_{[Nt]}^N = Y_{[Nt]}^N.$$

Puisque $n = [Nt]$, on a alors, en posant $\Delta t = \frac{1}{N}$,

$$\begin{aligned} \mathbb{E}(Z_{t+\Delta t}^N | Z_t^n = x) &= x \quad \text{et donc} \quad \mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^n = x) = 0 \\ \text{Var}(Z_{t+\Delta t}^N - Z_t^n | Z_t^n = x) &= \frac{1}{N} x(1-x) = x(1-x) \Delta t. \end{aligned}$$

Dans cette échelle de temps, la variance du processus n'explose pas quand N tend vers l'infini. On peut également remarquer que les sauts du processus $(Z_t^N, t \geq 0)$ sont d'amplitude $\frac{1}{N}$, et donc tendent vers 0 quand N tend vers l'infini. Par ailleurs le processus $(Z_t^N, t \geq 0)$ est un processus de Markov. Ainsi si ce processus a une limite quand $N \rightarrow \infty$, cette limite sera la diffusion définie comme solution de l'équation différentielle stochastique

$$dZ_t = \sqrt{Z_t(1-Z_t)} dB_t ; Z_0 = z. \quad (4.5.13)$$

Ce résultat peut être prouvé rigoureusement.

Remarquons tout d'abord que nécessairement, pour tout t , $Z_t \in [0, 1]$ comme limite d'une suite de nombres réels de $[0, 1]$. On peut alors montrer la convergence en loi du processus $(Z_t^N, t \geq 0)$, en tant que processus à trajectoires continues à droite et limitée à gauche sur tout intervalle de temps $[0, T]$ vers $(Z_t, t \geq 0)$. La preuve se fait en deux temps :

- Montrer que la suite Z^N admet une valeur d'adhérence (résultat difficile qui dépasse le cadre du cours) ;
- Montrer que cette valeur d'adhérence est unique. Il faut donc pour cela montrer qu'il y a unicité dans l'équation (4.5.13). Pour cela, on pourra appliquer le Théorème 2.3.8 en utilisant le fait que le processus reste borné. L'unicité entraîne en particulier que si Z atteint les bords, il y reste, c'est-à-dire que les points 0 et 1 sont absorbants.

4.5.2 Diffusion de Fisher-Wright avec mutation et sélection

Reprenons les calculs précédents, dans le cas d'un modèle de Wright-Fisher avec les paramètres de mutation introduits précédemment. Alors si $X_n^N = i$, la probabilité p_i est

définie par (4.4.12). On suppose qu'au moins un des paramètres α_1, α_2 est non nul, et on va comme précédemment étudier le processus

$$Z_t^N = \frac{1}{N} X_{[Nt]}^N.$$

Alors, les calculs donnent

$$\begin{aligned} \mathbb{E}(Z_{t+\Delta t}^N | Z_t^n = x) &= p_{Nx} \\ \mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^n = x) &= N(-\alpha_1 x + (1-x)\alpha_2) \Delta t \\ \mathbb{E}((Z_{t+\Delta t}^N - Z_t^n)^2 | Z_t^n = x) &= \mathbb{E}((Z_{t+\Delta t}^N)^2 | Z_t^n = x) - 2x \mathbb{E}(Z_{t+\Delta t}^N | Z_t^n = x) + x^2 \\ &= \frac{1}{N} p_{Nx} (1 - p_{Nx}) + p_{Nx}^2 - 2x p_{Nx} + x^2 \\ &= (x(1-x) + N((\alpha_1^2 + \alpha_2^2 + 2\alpha_1\alpha_2)x^2 - 2\alpha_2(\alpha_1 + \alpha_2)x + \alpha_2^2)) \Delta t. \end{aligned}$$

Pour que l'espérance conditionnelle de l'écart $Z_{t+\Delta t}^N - Z_t^n$ ait une chance de converger, nous allons nous placer sous l'hypothèse des mutations rares

$$\alpha_1 = \frac{\beta_1}{N}; \quad \alpha_2 = \frac{\beta_2}{N}. \quad (4.5.14)$$

Cela veut dire que les taux de mutation sont inversement proportionnels à la taille de la population. On a alors la proposition suivante

Proposition 4.5.1 *Sous l'hypothèse des mutations rares (4.5.14), et si $\frac{1}{N} X_0^N$ a une limite $z \in [0, 1]$ quand $N \rightarrow \infty$, le processus $(\frac{1}{N} X_{[Nt]}^N, t \geq 0)$ converge vers la solution de l'équation différentielle stochastique*

$$dZ_t = \sqrt{Z_t(1-Z_t)} dB_t + (-\beta_1 Z_t + (1-Z_t)\beta_2) dt; \quad Z_0 = z. \quad (4.5.15)$$

L'unicité est obtenue comme ci-dessus grâce au Théorème 2.3.8. Là-encore, $Z_t \in [0, 1]$.

Preuve. Sous l'hypothèse (4.5.14), $\mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^n = x)$ converge vers $(-\beta_1 x + (1-x)\beta_2)\Delta t$ quand $N \rightarrow \infty$. De plus, $\mathbb{E}((Z_{t+\Delta t}^N - Z_t^n)^2 | Z_t^n = x)$ converge alors vers $x(1-x)\Delta t$. Par des arguments analogues à ceux du paragraphe précédent, on en déduit le résultat. \square

Nous allons maintenant étudier le cas d'un modèle de sélection et nous allons faire une hypothèse de sélection rare, au sens où

$$s = \frac{r}{N}. \quad (4.5.16)$$

On considère toujours le processus

$$Z_t^N = \frac{1}{N} X_{[Nt]}^N,$$

défini à partir du modèle de Wright-Fisher avec sélection, c'est-à-dire quand la probabilité de transition de X^N est définie à partir des probabilités (4.4.10). On a alors la proposition suivante.

Proposition 4.5.2 *Sous l'hypothèse de sélection rare (4.5.16), et si $\frac{1}{N}X_0^N$ a une limite $z \in [0, 1]$ quand $N \rightarrow \infty$, le processus $(\frac{1}{N}X_{[Nt]}^N, t \geq 0)$ converge vers la solution de l'équation différentielle stochastique*

$$dZ_t = \sqrt{Z_t(1-Z_t)}dB_t + rZ_t(1-Z_t)dt ; Z_0 = z. \quad (4.5.17)$$

L'unicité est obtenue comme ci-dessus grâce au Théorème 2.3.8. Là-encore, $Z_t \in [0, 1]$.

Preuve. On a

$$\begin{aligned} \mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^N = x) &= N \left(\frac{-x^2s + sx}{sx + 1} \right) \Delta t \\ \mathbb{E}((Z_{t+\Delta t}^N - Z_t^n)^2 | Z_t^n = x) &= \left(x(1-x) + N \left(\frac{s^2x^2(1+2x+x^2)}{(xs+1)^2} \right) \right) \Delta t. \end{aligned}$$

Ainsi, en supposant (4.5.16), on obtient que $\mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^N = x)$ converge vers $rx(1-x)\Delta t$ quand $N \rightarrow \infty$. De plus, $\mathbb{E}((Z_{t+\Delta t}^N - Z_t^n)^2 | Z_t^n = x)$ converge alors vers $x(1-x)\Delta t$. Par des arguments analogues à ceux du paragraphe précédent, on en déduit le résultat. \square

Proposition 4.5.3 *Considérons maintenant le modèle le plus général avec mutation et sélection, où les probabilités de transition de $(X_n^N, n \in \mathbb{N})$ sont obtenues avec (4.4.12). Plaçons-nous sous les hypothèses de mutations et sélection rares (4.5.14) et (4.5.16). Alors, si $\frac{1}{N}X_0^N$ a une limite $z \in [0, 1]$, le processus $(Z_t^N, t \geq 0)$ défini par $Z_t^N = \frac{1}{N}X_{[Nt]}^N$, converge quand N tend vers l'infini vers la solution de l'équation différentielle stochastique*

$$dZ_t = \sqrt{Z_t(1-Z_t)}dB_t + rZ_t(1-Z_t)dt + (-\beta_1Z_t + (1-Z_t)\beta_2)dt ; Z_0 = z, \quad (4.5.18)$$

dès que $\frac{1}{N}X_0^N$ a pour limite $z \in [0, 1]$.

Remarque 4.5.4 *Remarquons que si l'on veut calculer des probabilités de fixation ou d'extinction d'un allèle, les calculs sont presque impossibles, car trop compliqués sur les modèles discrets. Ils deviennent accessibles avec les outils du calcul stochastique, comme nous l'avons vu dans le Chapitre 2.*

4.5.3 Autre changement d'échelle de temps

Reprenons le modèle précédent mais considérons maintenant le processus

$$R_t^N = \frac{1}{N}X_{[N^\gamma t]}^N,$$

où $0 < \gamma < 1$. Cela revient à dire que nous accélérons un peu moins le temps. Alors, en reprenant les calculs développés dans le modèle de mutation ci-dessus, et en supposant maintenant que

$$\alpha_1 = \frac{\beta_1}{N^\gamma} ; \alpha_2 = \frac{\beta_2}{N^\gamma},$$

nous obtiendrons que $\mathbb{E}((Z_{t+\Delta t}^N - Z_t^n)^2 | Z_t^n = x)$ tend vers 0 quand N tend vers l'infini et que $\mathbb{E}(Z_{t+\Delta t}^N - Z_t^n | Z_t^n = x)$ converge vers $(-\beta_1 x + (1-x)\beta_2)\Delta t$. On peut alors montrer que si $\frac{1}{N}X_0^N$ a pour limite y_0 quand $N \rightarrow \infty$, alors le processus $(R_t^N, t \geq 0)$ converge vers la fonction $(y(t), t \geq 0)$ (déterministe) solution de l'équation différentielle

$$dy(t) = (-\beta_1 y(t) + (1 - y(t))\beta_2)dt ; y(0) = y_0.$$

De même, dans le modèle de sélection avec

$$s = \frac{r}{N^\gamma},$$

on peut montrer que le processus $(R_t^N, t \geq 0)$ converge vers la fonction $(y(t), t \geq 0)$ (déterministe) solution de l'équation différentielle

$$dy(t) = ry(t)(1 - y(t))dt ; y(0) = y_0. \quad (4.5.19)$$

Dans l'approximation déterministe (4.5.19), il est facile de calculer l'équilibre. On voit que l'on atteint la fixation ou l'extinction selon que $r > 0$ ou $r < 0$.

Nous remarquons donc que si $0 < \gamma < 1$, la limite est déterministe et si $\gamma = 1$, la limite devient stochastique. Cette stochasticité provient de l'extrême variabilité due à un très grand nombre d'événements de reproduction d'espérance nulle (nous sommes dans une limite de type "Théorème de la limite centrale"). Bien-sûr, si $\gamma > 1$, les limites explosent, l'échelle de temps n'est plus adaptée à la taille de la population.

Il est très important de garder en tête ces relations entre les échelles de taille et de temps, et cette dichotomie entre un comportement déterministe qui prend en compte uniquement l'aspect macroscopique du phénomène et un comportement stochastique qui va en outre prendre en compte la très grande variabilité individuelle du système.

4.6 La coalescence : description des généalogies

Imaginons une généalogie d'individus diploïdes, à savoir les individus et toute leur descendance. Chacun de ces individus va avoir un certain nombre de descendants auxquels il aura transmis une des 2 copies de ses gènes à un locus donné. D'une génération sur l'autre, certains gènes ne seront pas transmis, mais d'autres pourront être transmis en plusieurs exemplaires. Il est naturel de chercher à savoir quelle est la généalogie d'un échantillon observé de gènes (ou d'individus) à une certaine génération. Notre but est de reconstruire l'histoire généalogique des gènes, selon les contraintes démographiques de la population observée et des possibles mutations, jusqu'à l'ancêtre commun le plus récent de ces gènes. Nous n'allons pas prendre en compte la totalité de la population, et nous nous concentrerons sur l'échantillon d'intérêt. C'est une approche rétrospective. Comme

précédemment, nous allons assimiler la population diploïde de taille M à une population haploïde de taille $2M = N$.

On appelle **lignage** l'ascendance d'un gène. Lorsque deux lignages se rejoignent chez un gène ancestral, on dit qu'ils **coalescent** ou qu'il s'est produit **un événement de coalescence**.

La théorie de la coalescence décrit donc simplement le processus de coalescence des gènes d'un échantillon depuis la génération présente jusqu'à leur ancêtre commun.

4.6.1 Asymptotique quand N tend vers l'infini : le coalescent de Kingman

Revenons au modèle de Wright-Fisher neutre, c'est-à-dire sans mutation ni sélection. Nous imaginons maintenant un petit échantillon d'individus observés dans une population de taille N très grande, que l'on assimilera à l'infini.

Nous avons vu que la probabilité $h(n)$ d'avoir hétérozygotie au temps n est donnée, quand la taille de la population est N , par la formule (4.4.8). Supposons maintenant que N soit très grand. Si n est petit, cette formule apporte peu d'information, hormis que $h(n)$ est de l'ordre de $h(0)$. Il est alors beaucoup plus intéressant d'étudier ce qui se passe en temps long (n grand). Nous allons donc faire l'hypothèse que nos observations sont obtenues en un temps n très long, de l'ordre de N . Alors, puisque $(1 - x) \approx e^{-x}$ quand x est proche de 0, on en déduit que

$$h(n) \approx e^{-\frac{n}{N}} h(0),$$

et ainsi que la probabilité d'hétérozygotie (appelée coefficient d'hétérozygotie en génétique des populations) décroît exponentiellement vite en la variable $\frac{n}{N}$.

Nous considérons à partir de maintenant un échantillon de k individus, issu de cette population de taille N très grande ($N \rightarrow \infty$). Calculons la probabilité pour que deux individus (au moins) aient le même parent. Cela aura lieu si l'un des trois événements suivants est satisfait :

- exactement deux individus parmi les k ont le même parent,
- 3 individus au moins ont le même parent,
- au moins deux paires d'individus ont un parent commun.

La probabilité des deux derniers événements est de l'ordre de $\frac{1}{N^2}$ et ces événements ont donc des probabilités négligeables par rapport à la probabilité du premier événement, qui vaut

$$\frac{k(k-1)}{2} \frac{1}{N}.$$

Ainsi, quand N tend vers l'infini, la probabilité pour que deux individus aient le même parent est de l'ordre de $\frac{k(k-1)}{2} \frac{1}{N}$.

En raisonnant comme dans la preuve de (4.4.8), nous obtenons alors que la probabilité qu'il n'y ait pas eu de parents communs pour ces k individus dans les n premières générations est, quand N tend vers l'infini et n est de l'ordre de N , proportionnelle à

$$\left(1 - \frac{k(k-1)}{2} \frac{1}{N}\right)^n \sim \exp\left(-\frac{k(k-1)}{2} \frac{n}{N}\right). \quad (4.6.20)$$

A partir de maintenant, nous considérons les k individus de l'échantillon à un certain instant, et remontons le temps. Définissons T , le premier instant où les k individus ont un ancêtre commun dans leur généalogie. Nous déduisons de (4.6.20) que

$$P(T > n) \sim \exp\left(-\frac{k(k-1)}{2} \frac{n}{N}\right),$$

quand N tend vers l'infini et $\frac{n}{N}$ tend vers une constante.

L'analogie avec le comportement d'une variable exponentielle va nous faire comprendre dans quelle échelle de temps nous devons nous placer.

En effet, rappelons que si τ suit une loi exponentielle de paramètre λ , alors

$$\mathbb{P}(\tau > t) = e^{-\lambda t}.$$

Nous allons changer de temps et considérer comme nouvelle unité de temps N générations, en posant $t \approx \frac{n}{N}$, c'est à dire que si t est la nouvelle unité de temps, on va remplacer n par $[Nt]$ (qui désigne la partie entière de Nt). Ainsi donc, si l'on considère k individus, **le temps de coalescence, c'est-à-dire d'apparition du premier ancêtre commun à deux individus dans le passé, se comporte comme une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$ et donc de moyenne $\frac{2}{k(k-1)}$.**

Rappel (Voir MAP 311) : Soient τ_1, \dots, τ_m , m variables aléatoires exponentielles indépendantes de paramètre λ . Alors la variable aléatoire $\inf_{l=1, \dots, m} T_l$ suit une loi exponentielle de paramètre $m\lambda$.

Remarque 4.6.1 Le temps de coalescence de 2 lignées est une variable aléatoire exponentielle de paramètre 1 et si l'on a k individus, alors le premier temps auquel une paire a un ancêtre commun est donc l'infimum de $\frac{k(k-1)}{2}$ variables aléatoires exponentielles de paramètre 1 indépendantes.

On va alors définir le k -coalescent comme limite quand N tend vers l'infini du processus qui décrit la généalogie de l'échantillon de k individus au sens suivant.

Définition 4.6.2 Soit $k \in \mathbb{N}^*$. On appelle *k-coalescent* la chaîne de Markov $(\Pi_t)_t$ à valeurs dans l'ensemble \mathcal{P}_k des partitions de $\{1, \dots, k\}$ définie de la manière suivante :

- $\Pi_0 = \{\{1\}, \dots, \{k\}\}$.
- Soit T_i le i -ème temps de coalescence. On pose $T_0 = 0$. Alors les intervalles de temps $T_i - T_{i-1}$ sont indépendants et suivent des lois exponentielles de paramètre $\frac{(k-i)(k-i+1)}{2}$.
- A chaque temps de saut, deux blocs de la partition sont choisis uniformément parmi les paires de blocs existantes et coalescent, au sens où les deux sous blocs sont regroupés en un seul.

Bien-sûr, vu la définition, il existe un temps T où l'on a trouvé le plus récent ancêtre commun aux k individus. A chaque temps de coalescence, le nombre d'éléments de la partition diminue de 1, et donc il sera réduit à un élément au bout de k événements de coalescence, et $T = T_{k-1}$.

Remarquons que la définition donnée du coalescent permet d'en déduire facilement un algorithme de simulation :

- On se donne k individus. On pose $\tilde{\Pi}_0 = \{\{1\}, \dots, \{k\}\}$.
- On simule une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$. Pour ce faire, on considère une variable aléatoire U de loi uniforme sur $[0, 1]$ et on pose $T_1 = \frac{2}{k(k-1)} \log(1/U)$. (Voir MAP 311).
- On choisit uniformément au hasard deux blocs distincts de la partition, (donc avec probabilité $\frac{2}{k(k-1)}$). On regroupe ces deux blocs en un seul bloc. On appelle alors $\tilde{\Pi}_1$ cette nouvelle partition.
- On réitère cette procédure. Après l'étape $i-1$, on simule une variable aléatoire exponentielle de paramètre $\frac{(k-i+1)(k-i)}{2}$. On choisit uniformément au hasard deux blocs distincts de la partition composée de $k-i+1$ blocs, (donc avec probabilité $\frac{2}{(k-i)(k-i+1)}$). On regroupe ces blocs en un seul ensemble. On obtient ainsi $\tilde{\Pi}_i$.
- On pose alors

$$\Pi_t = \sum_i \tilde{\Pi}_i \mathbf{1}_{\{T_i \leq t < T_{i+1}\}}. \quad (4.6.21)$$

Théorème 4.6.3 Soit T le temps du plus récent ancêtre commun (PRAC) des k individus de l'échantillon. Alors

$$\mathbb{E}(T) = 2 \left(1 - \frac{1}{k}\right). \quad (4.6.22)$$

Preuve. On écrit que

$$T = T_{k-1} = T_{k-1} - T_{k-2} + \dots + T_1.$$

T est donc la somme de $k-1$ variables aléatoires exponentielles indépendantes \mathcal{E}_i de paramètres $\frac{(k-i+1)(k-i)}{2}$. Ainsi,

$$\mathbb{E}(T) = \sum_{i=1}^{k-1} \frac{2}{(k-i)(k-i+1)} = 2 \sum_{i=1}^{k-1} \left(\frac{1}{k-i} - \frac{1}{k-i+1} \right) = 2 \left(1 - \frac{1}{k}\right).$$

□

Théorème 4.6.4 (cf. Durrett [6]). Si π est une partition de $\{1, \dots, k\}$ en l blocs, soit $\pi = (B_1, B_2, \dots, B_l)$. Alors cette partition pourra être réalisée au temps T_{k-l} , et

$$\mathbb{P}(\Pi_{T_{k-l}} = \pi) = \frac{l! (k-l)!(l-1)!}{k! (k-1)!} \prod_{i=1}^l (\text{Card}(B_i))!, \quad (4.6.23)$$

où l'on rappelle que $k! = k(k-1)\dots 1$. Nous avons ainsi une description de la loi du k -coalescent.

Dans la suite, nous noterons

$$c_{k,l} = \frac{l! (k-l)!(l-1)!}{k! (k-1)!} \quad (4.6.24)$$

et

$$w(\pi) = \prod_{i=1}^l (\text{Card}(B_i))! \quad (4.6.25)$$

si B_1, \dots, B_l sont les blocs de la partition π . Ainsi la probabilité qu'au $(k-l)$ -ième temps de coalescence, le k -coalescent à l blocs soit égal à π est le produit du terme $c_{k,l}$ qui ne dépend que de k et l , et du poids $\prod_{i=1}^l (\text{Card}(B_i))!$ qui favorise les partitions inégales. (On vérifiera à titre d'exemple que si on a deux partitions de 5 individus, l'une de 2 et 3 individus et l'autre de 4 et 1 individus, la première donnera le poids 12 et la deuxième le poids 24).

Remarquons que le temps T_{k-l} est indépendant de l'état de la partition.

Preuve. On procède par induction descendante sur l . Quand $l = k$, la seule partition π possible est la partition composée des blocs singletons, et la probabilité de réalisation de π est 1, ce qui est également donné par le membre de droite de (4.6.23) pour $l = k$.

Supposons que (4.6.23) soit vraie pour toute partition de taille l et considérons une partition η de taille $l-1$. On notera $\xi < \eta$ si $\text{Card}(\xi) = \text{Card}(\eta) + 1$ et si η est obtenue après regroupement de deux blocs de ξ . Quand $\xi < \eta$, il y a exactement un événement de coalescence qui va faire passer de ξ à η . Alors,

$$\mathbb{P}(\Pi_{T_{k-l+1}} = \eta | \Pi_{T_{k-l}} = \xi) = \begin{cases} \frac{2}{l(l-1)} & \text{si } \xi < \eta \\ 0 & \text{sinon.} \end{cases} \quad (4.6.26)$$

On a donc

$$\mathbb{P}(\Pi_{T_{k-l+1}} = \eta) = \frac{2}{l(l-1)} \sum_{\xi < \eta} \mathbb{P}(\Pi_{T_{k-l}} = \xi). \quad (4.6.27)$$

On pourra décrire toutes les partitions ξ possibles de la manière suivante.

Etant donnée une partition ξ de $\{1, \dots, k\}$ en l blocs, il est naturel d'ordonner ses blocs en $\xi_1, \xi_2, \dots, \xi_l$, où ξ_1 est le bloc contenant 1, ξ_2 celui contenant le plus petit nombre qui n'est pas dans ξ_1, \dots

Si les entiers $\lambda_1, \dots, \lambda_{l-1}$ sont les tailles des blocs de la partition η , alors pour j tel que $1 \leq j \leq l-1$, et m tel que $1 \leq m < \lambda_j$, on associera une partition $\xi < \eta$, où les blocs de ξ auront les tailles $\lambda_1, \dots, \lambda_{j-1}, m, \lambda_j - m, \lambda_{j+1}, \dots, \lambda_{l-1}$. Remarquons qu'il y a $\frac{1}{2} \binom{\lambda_j}{m}$ choix d'une telle partition $\xi < \eta$, où l'on a coupé le j -ème bloc en $\lambda_j - m$ et m individus.

Utilisant l'hypothèse de récurrence, nous obtenons alors que

$$\begin{aligned} \mathbb{P}(\Pi_{T_{k-l+1}} = \eta) &= \frac{2}{l(l-1)} \sum_{j=1}^{l-1} \sum_{m=1}^{\lambda_j-1} \frac{1}{2} \binom{\lambda_j}{m} c_{k,l} \lambda_1! \dots \lambda_{j-1}! m! (\lambda_j - m)! \lambda_{j+1}! \dots \lambda_l! \\ &= w(\eta) \frac{c_{k,l}}{l(l-1)} \sum_{j=1}^{l-1} \sum_{m=1}^{\lambda_j-1} 1. \end{aligned} \quad (4.6.28)$$

La double somme vaut

$$\sum_{j=1}^{l-1} (\lambda_j - 1) = k - (l - 1).$$

On vérifie alors facilement que

$$\frac{c_{k,l}}{l(l-1)} (k - l + 1) = c_{k,l-1}.$$

Le résultat est donc prouvé par induction. □

Théorème 4.6.5 *Soit σ une permutation choisie au hasard parmi les permutations de $\{1, \dots, l\}$. Soit $\lambda_j = \text{Card}(\xi_{\sigma(j)})$ la taille du j -ième bloc de la partition $\Pi_{T_{k-l}}$, quand les blocs sont réarrangés suivant σ et que l'on est parti de la numérotation naturelle. Alors $(\lambda_1, \dots, \lambda_l)$ suit une loi uniforme sur les vecteurs de $(\mathbb{N}^*)^l$ de somme k .*

Par exemple, pour $l = 2$, si on tire au hasard un des deux blocs de $\Pi_{T_{k-2}}$, la loi de sa taille est uniformément distribuée sur $\{1, \dots, k-1\}$.

Preuve. Chaque réarrangement ordonné des l blocs de la partition $\Pi_{T_{k-l}} = \Pi$ (constituée de k individus) a la probabilité $\frac{c_{k,l} w(\Pi)}{l!}$ d'être choisi. Si l'on garde seulement l'information

sur les tailles des blocs, alors la probabilité que les blocs aient des tailles égales au vecteur $(\lambda_1, \dots, \lambda_l)$, est égale (par la formule des probabilités conditionnelles) à

$$\frac{c_{k,l}w(\Pi)}{l!} \frac{k!}{\lambda_1! \lambda_2! \dots \lambda_l!} = \frac{(k-l)!(l-1)!}{(k-1)!} = \frac{1}{\binom{k-1}{l-1}}.$$

La quantité finale ne dépend que de k et de l , et pas du vecteur $(\lambda_1, \dots, \lambda_l)$. Ainsi, la distribution est uniforme. Vérifions que le dénominateur de la dernière fraction donne le nombre de vecteurs d'entiers positifs de somme égale à k . On imagine k boules séparées en l groupes par $l-1$ morceaux de carton. Par exemple, si $k=10$ et $l=4$, on pourrait avoir

$$OOO|O|OOOO|OO.$$

Les $l-1$ morceaux de carton peuvent se déplacer dans les $k-1$ espaces entre les boules. Il y a donc $\binom{k-1}{l-1}$ choix possibles de vecteurs de taille l d'entiers positifs de somme égale à k . \square

Comme conséquence de ce théorème, on a le résultat étonnant suivant.

Théorème 4.6.6 *La probabilité que le plus récent ancêtre commun d'un groupe de k individus soit le même que celui de la population totale converge vers $\frac{k-1}{k+1}$ quand la taille N de la population tend vers l'infini.*

Par exemple, quand $k=2$, cette probabilité vaut $\frac{1}{3}$ et quand $k=10$, elle est de $\frac{9}{11}$.

Preuve. Soit N la taille de la population. Les k individus n'ont pas le même PRAC (plus récent ancêtre commun) que la population totale de taille N , si et seulement si au premier branchement (dans le sens du temps physique), dans l'arbre de coalescence de la population, on a une partition en deux sous-arbres, et les k individus de l'échantillon doivent nécessairement appartenir à l'un ou à l'autre de ces sous-blocs. On notera $PRAC_N \neq PRAC_k$ l'événement "le PRAC de la population totale de taille N diffère de celui du groupe d'individu de taille K ". Notons $\mathcal{P}(i, N-i)$ l'événement "Au premier branchement, on a partition en 2 blocs de taille $(i, N-i)$ ". On a alors

$$\mathbb{P}(PRAC_N \neq PRAC_k) = \sum_{i=1}^N \mathbb{P}(PRAC_N \neq PRAC_k | \mathcal{P}(i, N-i)) \times \mathbb{P}(\mathcal{P}(i, N-i)) \quad (4.6.29)$$

La probabilité d'existence de la partition $(i, N-i)$ est uniforme, donnée par le théorème 4.6.5 pour $l=2$, et vaut donc $\frac{1}{N-1}$. Sachant $\mathcal{P}(i, N-i)$, l'événement $PRAC_N \neq PRAC_k$ sera réalisé si les k individus sont soit dans le bloc de taille i , soit dans le bloc de taille $N-i$. Comme N est très grand, on peut supposer que l'échantillonnage des k individus parmi les N se fait suivant un tirage avec remise. Alors on aura

$$\mathbb{P}(PRAC_N \neq PRAC_k | \mathcal{P}(i, N-i)) \sim \left(\frac{i}{N}\right)^k + \left(\frac{N-i}{N}\right)^k.$$

Ainsi, on en déduit finalement que

$$\mathbb{P}(PRAC_N \neq PRAC_k) \sim \frac{1}{N-1} \sum_{i=1}^k \left(\left(\frac{i}{N} \right)^k + \left(\frac{N-i}{N} \right)^k \right). \quad (4.6.30)$$

En faisant tendre N vers l'infini, et en utilisant la convergence des sommes de Riemann, on obtient alors à la limite que

$$\mathbb{P}(PRAC_N \neq PRAC_k) \sim \int_0^1 (x^k + (1-x)^k) dx = \frac{2}{k+1}.$$

Puisque $\mathbb{P}(PRAC_N = PRAC_k) = 1 - \mathbb{P}(PRAC_N \neq PRAC_k) = 1 - \frac{2}{k+1}$, on en déduit le résultat. \square

4.6.2 Mutation sur le coalescent

Le modèle précédent est un modèle très simple, car sans mutation (mais peut être justifié dans certaines échelles de temps). En fait de nombreuses mutations se produisent, à chaque reproduction. On suppose que des mutations arrivent sur l'arbre et qu'il y a tant d'allèles possibles que chaque mutation est toujours d'un type jamais vu avant. En effet, si par exemple, un gène consiste en 500 nucléotides (les bases A, T, G ou C), le nombre de séquences d'ADN possibles est $4^{500} = 10^{301}$, et il est raisonnable de supposer que la probabilité d'avoir deux mutations au même nucléotide en un temps raisonnable est négligeable.

Il y a deux manières de prendre en compte ces mutations : soit on compte le nombre total de mutations (ISM, infinite site model, modèle de Kimura et Crow), soit on compte le nombre d'allèles différents que l'on obtient dans un échantillon de taille k (IAM, infinite allele model). La première est plus simple du point de vue mathématique, mais la deuxième correspond à des observations biologiques. Comme précédemment, k désigne la taille de l'échantillon que l'on considère. On appelle M_k le nombre de mutations et A_k le nombre d'allèles différents.

Exemple :

Dans la figure 4.2, on a $M_8 = 10$, $A_8 = 7$.

Il est très intéressant de connaître les lois de ces variables aléatoires.

Par exemple, reprenons les exemples donnés dans le livre de Durrett (cf. [6]) : dans une étude de drosophiles, portant sur 60 gènes à un certain locus, on a répertorié 18 allèles uniques, 3 allèles ayant 2 représentations, 1 en ayant 4 et 1 en ayant 32. Ainsi,

$$a_1 = 18, a_2 = 3, a_4 = 1, a_{32} = 1,$$

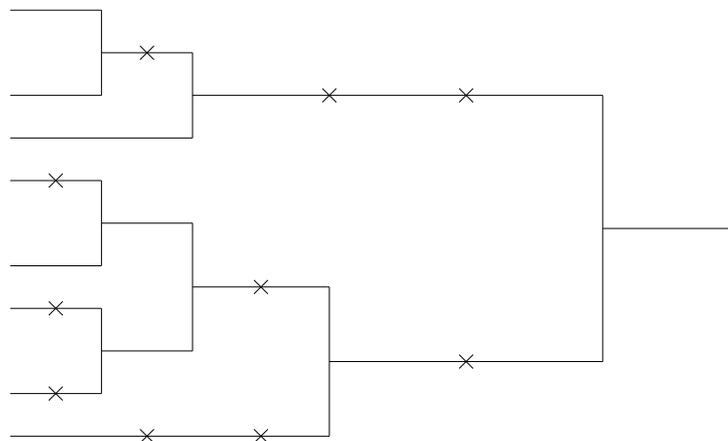


FIG. 4.2 – Mutations sur le coalescent

où a_i est le nombre d'allèles possédant i représentants parmi les mouches de l'échantillon. De même l'étude de 146 gènes d'une autre variété de drosophile à ce même locus donne

$$a_1 = 20, a_2 = 3, a_5 = 2, a_6 = 2, a_8 = 1, a_{11} = 1, a_{68} = 1.$$

On peut donc se demander comment évolue la loi de la variable aléatoire A_k , et c'est à cela que répondra la célèbre formule d'échantillonnage d'Ewens.

Essayons tout d'abord de comprendre comment l'on construit un coalescent avec mutations.

4.6.3 Le coalescent avec mutation

Revenons ici au modèle de Wright-Fisher pour une population de taille N dans laquelle nous considérons un échantillon de taille k . Comme précédemment, les individus de la génération n choisissent leur parent dans la génération $n - 1$, uniformément parmi les N individus, indépendamment les uns des autres. On a vu que la probabilité d'avoir un événement de coalescence vaut $\frac{k(k-1)}{2} \frac{1}{N}$. Nous supposons de plus que l'allèle du descendant n'est pas automatiquement héritable de celui du parent et que l'on peut avoir mutation avec probabilité μ . Cela veut dire que l'allèle d'un individu est hérité de celui du parent avec probabilité $1 - \mu$ et qu'avec probabilité μ , on a un nouveau type. Nous allons nous intéresser au modèle limite, dans l'asymptotique du coalescent de Kingman.

Comme précédemment, nous obtenons que la probabilité qu'il n'y ait pas eu de parents communs pour les k individus de l'échantillon choisi est, quand N tend vers l'infini, et si $\frac{n}{N}$ tend vers une constante, de l'ordre de

$$\exp\left(-\frac{k(k-1)}{2} \frac{n}{N}\right).$$

Par ailleurs, si l'on suppose que $\mu \sim \frac{\theta}{2N}$, la probabilité pour qu'il y ait eu une mutation à la génération précédente est de l'ordre de $k\frac{\theta}{2N}$, qui est la probabilité qu'il y ait eu exactement une mutation. En effet, la probabilité qu'il y ait au moins deux mutations est d'ordre $\frac{1}{N^2}$. Ainsi, nous obtenons que la probabilité qu'il n'y ait pas eu de mutation au cours des n premières générations et pour les k individus est de l'ordre de

$$\left(1 - k\frac{\theta}{2N}\right)^n \approx \exp\left(-k\frac{\theta}{2} \frac{n}{N}\right).$$

Donc, si S désigne le premier instant où au moins un individu parmi les k individus de l'échantillon provient d'une mutation, on a

$$\mathbb{P}(S > n) \approx \exp\left(-k\frac{\theta}{2} \frac{n}{N}\right).$$

L'on voit donc que le changement d'échelle de temps $n = [Nt]$ qui permet d'obtenir le coalescent de Kingman entraîne alors que dans cette échelle rapide et si $\mu \sim \frac{\theta}{2N}$, le temps d'apparition d'une mutation suit une loi exponentielle de paramètre $k\frac{\theta}{2}$, et donc de moyenne $\frac{2}{k\theta}$.

On obtient un modèle de coalescent avec taille infinie et taux de mutation $\frac{\theta}{2}$. On a donc sur chaque branche un taux $\frac{\theta}{2}$ de mutation et sur chaque couple de branches un taux 1 de coalescence. La construction algorithmique du coalescent avec mutation se fait exactement comme celle de la définition 2.6.2. Toutefois, quand on suit l'arbre de coalescence, sur chaque branche on regarde une cloche exponentielle de paramètre $\frac{\theta}{2}$. Si elle sonne on remplace l'allèle du parent par un nouveau type qui va se transmettre ensuite (choisi dans un océan de types différents). Plus précisément, entre les instants de coalescence T_{i-1} et T_i , on a i individus et le temps de mutation suivra une loi exponentielle de paramètre $i\frac{\theta}{2}$.

Remarque fondamentale : Si nous remontons le temps dans le coalescent de Kingman, à partir d'un échantillon de taille $k+1$, on sait qu'un événement de coalescence aura lieu au taux $\frac{k(k+1)}{2}$ et une mutation au taux $(k+1)\frac{\theta}{2}$. Alors, si on tire au hasard une branche, on va rencontrer une mutation avant un événement de coalescence avec probabilité

$$\begin{aligned} \mathbb{P}(S < T) &= \int_0^\infty (k+1)\frac{\theta}{2} e^{-(k+1)\frac{\theta}{2}u} \int_u^\infty \frac{k(k+1)}{2} e^{-\frac{k(k+1)}{2}v} dv du \\ &= \int_0^\infty (k+1)\frac{\theta}{2} e^{-(k+1)\frac{\theta}{2}u} e^{-\frac{k(k+1)}{2}u} du = \frac{\theta}{\theta+k}. \end{aligned}$$

De même, on va rencontrer une coalescence avec probabilité

$$\frac{k}{\theta + k}.$$

4.6.4 Urne de Hoppe, restaurant chinois et modèle de Hubbell

Nous allons présenter ici trois modèles équivalents à notre coalescent avec mutation.

- **L'urne de Hoppe.** On considère une urne qui va contenir une boule noire de masse θ et des boules colorées de masse 1. Au démarrage, il y a une boule noire. Au k -ième tirage, on tire une boule (avec remise) au prorata de sa masse. Si une boule colorée est tirée, on la remet dans l'urne en rajoutant une autre boule de même couleur. Si la boule est noire, on rajoute une boule d'une nouvelle couleur. On suppose que le nombre de couleurs est infini. Au k -ième tirage, il y a donc k boules dans l'urne. Le choix de la boule noire correspond à une nouvelle mutation et le choix d'une boule colorée à un événement de coalescence. On réitère récursivement le processus. Quand on remonte le temps du tirage $k+1$ au tirage k dans ce modèle, on trouve une mutation avec probabilité $\frac{\theta}{\theta+k}$ et un événement de coalescence avec probabilité $\frac{k}{\theta+k}$. A_k est ici le nombre de couleurs différentes après k tirages.

On a donc le théorème suivant.

Théorème 4.6.7 *La généalogie de k particules dans un coalescent avec mutation peut-être décrite (du point de vue de la loi) en simulant le processus de tirage et remplacement de couleurs dans une urne de Hoppe après k étapes.*

Remarque : cette construction est très pratique pour la simulation d'un coalescent avec mutation.

- **Le modèle du restaurant chinois.** On considère une infinité de tables. Supposons k convives assis à un certain nombre de tables. Un nouveau convive arrive. Avec probabilité $\frac{\theta}{\theta+k}$, il s'assied à une nouvelle table et avec probabilité $\frac{1}{\theta+k}$, il choisit un convive au hasard et s'assoit à son côté. Le nombre d'allèles (d'espèces) distincts a même loi que le nombre de table occupées.

4.6.5 Le modèle écologique de Hubbell

L'approximation conduisant au coalescent avec mutation a aussi son intérêt en écologie, à travers le modèle écologique de Hubbell. Décrivons-le rapidement.

Le modèle écologique de Hubbell (cf. Hubbell). C'est un modèle d'écologie quantitative. On considère une forêt comprenant N sites occupés par N arbres d'espèces variées. Chaque arbre qui meurt laisse place à un nouvel arbre dont l'espèce est soit ré-échantillonnée parmi les $k-1$ arbres restants avec probabilité $1-\mu$, soit est remplacée par une nouvelle espèce avec probabilité μ . On suppose que les espèces migrantes proviennent d'une métacommunauté d'arbres composée d'un nombre infini d'espèces. La probabilité μ s'appelle le taux de spéciation.

4.6.6 Loi du nombre d'allèles distincts, formule d'Ewens

Étudions la loi de A_k . Remarquons que dans le modèle du restaurant chinois, c'est le nombre de tables occupées par les k premiers convives.

Théorème 4.6.8 1) La variable aléatoire A_k peut s'écrire

$$A_k = \sum_{i=1}^k \varepsilon_i,$$

où les variables aléatoires ε_i sont des variables aléatoires de Bernoulli indépendantes de loi

$$\mathbb{P}(\varepsilon_i = 1) = \frac{\theta}{\theta + i - 1}. \quad (4.6.31)$$

2) On en déduit que

$$\mathbb{E}(A_k) \sim_{k \rightarrow \infty} \theta \ln k ; \text{Var}(A_k) \sim_{k \rightarrow \infty} \theta \ln k. \quad (4.6.32)$$

3) Si Φ désigne la fonction de répartition de la loi normale centrée réduite, alors pour tout nombre réel x , on a

$$\mathbb{P}\left(\frac{A_k - \mathbb{E}(A_k)}{\sqrt{\text{Var}(A_k)}} \leq x\right) \rightarrow \Phi(x), \quad (4.6.33)$$

quand k tend vers l'infini.

Preuve. Dans les différents modèles, on pose $\varepsilon_i = 1$ si le i -ième allèle provient d'une mutation, (ou si le i -ième convive choisit une nouvelle table, ou si le i -ième tirage donne une nouvelle couleur). On pose $\varepsilon_i = 0$ sinon. Alors les ε_i sont des variables aléatoires de Bernoulli indépendantes et de paramètre $\frac{\theta}{\theta+i-1}$. Ainsi, nous avons donc

$$\mathbb{E}(A_k) = \sum_{i=1}^k \mathbb{E}(\varepsilon_i) = \sum_{i=1}^k \frac{\theta}{\theta + i - 1} ; \text{Var}(A_k) = \sum_{i=1}^k \frac{\theta(i-1)}{(\theta + i - 1)^2}.$$

En utilisant la décroissance de $x \rightarrow \frac{\theta}{\theta+x}$, on montre que

$$\int_0^k \frac{\theta}{\theta+x} dx \leq \mathbb{E}(A_k) \leq 1 + \int_0^{k-1} \frac{\theta}{\theta+x} dx.$$

On en déduit par intégration que

$$\mathbb{E}(A_k) \sim_{k \rightarrow \infty} \theta \ln(k).$$

On a de plus que

$$\text{Var}(A_k) - \mathbb{E}(A_k) = - \sum_{i=0}^{k-1} \left(\frac{\theta}{\theta + i} \right)^2,$$

somme partielle d'une série convergente. Et donc $\text{Var}(A_k)$ se comporte comme $\mathbb{E}(A_k)$. La troisième assertion découle d'un théorème de la limite centrale qui généralise le théorème central limite classique : ici, les variables aléatoires sont indépendantes mais pas de même loi. Le théorème est le suivant.

Théorème 4.6.9 *Soit $(X_i)_i$ une suite de variables indépendantes de carré intégrable centrées. Introduisons $S_n = X_1 + \dots + X_n$ et $s_n^2 = \mathbb{E}(S_n^2)$. Alors la condition de Lindenberg*

$$\forall \varepsilon > 0, \frac{1}{s_n^2} \sum_{i=1}^n \mathbb{E}(X_i^2 \mathbf{1}_{|X_i| > \varepsilon s_n}) \rightarrow_{n \rightarrow \infty} 0$$

est réalisée si et seulement si $\lim_n (s_n^{-2} \max_{1 \leq i \leq n} \mathbb{E}(X_i^2)) = 0$ et dans ce cas, la suite $(\frac{S_n}{s_n})_n$ converge en loi vers une variable aléatoire normale centrée réduite.

Appliquons ce théorème dans notre contexte. On peut montrer que

$$\max_{1 \leq i \leq n} \mathbb{E}(X_i^2) = \max_{1 \leq i \leq n} \frac{\theta(i-1)}{(\theta+i-1)^2} = \frac{\theta(n-1)}{(\theta+n-1)^2},$$

et nous savons que $s_n^2 \sim \theta \ln n$. Ainsi la condition $\lim_n (s_n^{-2} \max_{1 \leq i \leq n} \mathbb{E}(X_i^2)) = 0$ est réalisée et nous pouvons appliquer le théorème. Ainsi, la suite $\left(\frac{A_k - \mathbb{E}(A_k)}{\sqrt{\text{Var}(A_k)}} \right)_k$ converge en loi vers une loi normale centrée réduite. \square

Comme conséquence immédiate du théorème 4.6.8, nous voyons que la variable aléatoire $\frac{A_k}{\ln k}$ est un estimateur asymptotiquement normal du paramètre inconnu θ . Et nous savons de plus que sa variance tend très lentement vers 0, à la vitesse de $\frac{1}{\ln(k)}$. Cela donne une très mauvaise vitesse de convergence de l'estimateur, et nécessite donc de très grandes tailles d'échantillons. Ainsi, si l'on veut estimer θ avec une erreur de 0.1, on doit utiliser un échantillon de taille approximativement $k = e^{100}$. Il serait naturel de chercher un autre moyen d'estimer θ . Mais en fait il n'y en a pas d'autre aussi bon. (cf. Durrett)

Le dernier résultat de ce chapitre décrit le comportement asymptotique du nombre d'allèles. Il est dû à Ewens (cf. Ewens [8]), et donne la distribution complète dans l'échantillon, le nombre d'allèle présents et leur quantité. Pour $j \in \{1, \dots, k\}$, soit N_j le nombre d'allèles portés par j individus dans l'échantillon. Par exemple, N_1 est le nombre d'allèles singletons. Les N_j vérifient en particulier

$$\sum_{j=1}^k j N_j = k. \quad (4.6.34)$$

On a alors le théorème suivant.

Théorème 4.6.10 (*Ewens sampling formula*) *Considérons k entiers n_1, \dots, n_k inférieurs ou égaux à k avec $\sum_{j=1}^k jn_j = k$. Alors on a*

$$\mathbb{P}(N_1 = n_1, \dots, N_k = n_k) = \prod_{j=1}^k \frac{j}{j-1+\theta} \frac{\left(\frac{\theta}{j}\right)^{n_j}}{n_j!} = \frac{k!}{\theta_{(k)}} \prod_{j=1}^k \frac{\left(\frac{\theta}{j}\right)^{n_j}}{n_j!}, \quad (4.6.35)$$

où $\theta_{(k)} = \theta(\theta+1)\dots(\theta+k-1)$.

La formule peut paraître très compliquée mais semblera plus familière si on l'écrit

$$C_{n,\theta} \prod_{j=1}^n e^{-\frac{\theta}{j}} \frac{\left(\frac{\theta}{j}\right)^{n_j}}{n_j!},$$

où $C_{n,\theta}$ est une constante qui dépend de θ et n , telle que la somme des probabilités ci-dessus vaut 1. En d'autres mots, si on considère des variables aléatoires $Y_1, \dots, Y_j, \dots, Y_k$ indépendantes et de loi de Poisson de paramètres respectifs $\frac{\theta}{j}$, la partition allélique a même loi que la distribution de

$$\left(Y_1, \dots, Y_k \mid \sum_j jY_j = k \right).$$

Ce résultat sera justifié ultérieurement.

Preuve. Il suffit de montrer que la distribution des couleurs dans l'urne de Hoppe au temps k est donnée par la formule d'Ewens. Nous le montrons par récurrence sur k . Quand $k = 1$, la partition $N_1 = 1$ a probabilité 1 et le résultat est prouvé. Supposons maintenant que la propriété soit prouvée pour tout temps inférieur à $k - 1$. Supposons qu'au temps k , on ait la distribution $N_1 = n_1, \dots, N_k = n_k$. Notons $n = (n_1, \dots, n_k)$ et soit \bar{n} l'état de la répartition allélique au temps précédent. Notons

$$P_\theta(n) = \frac{k!}{\theta_{(k)}} \prod_{j=1}^k \frac{\left(\frac{\theta}{j}\right)^{n_j}}{n_j!}.$$

Nous allons montrer que

$$\sum_{\bar{n}} P_\theta(\bar{n})p(\bar{n}, n) = P_\theta(n). \quad (4.6.36)$$

On a plusieurs possibilités :

- Si $\bar{n}_1 = n_1 - 1$: cela veut dire qu'une nouvelle couleur vient d'être ajoutée. Dans ce cas, la probabilité de transition pour l'urne de Hoppe est de

$$p(\bar{n}, n) = \frac{\theta}{\theta + k - 1}.$$

En notant $\mathbf{N} = (N_1, \dots, N_k)$ on aura par ailleurs par la formule d'Ewens que

$$\frac{P_\theta(n)}{P_\theta(\bar{n})} = \frac{k}{\theta + k - 1} \frac{\theta}{n_1}.$$

- Supposons que pour $1 \leq j \leq k$, on a $\bar{n}_j = n_j + 1$ et que $\bar{n}_{j+1} = n_{j+1} - 1$. Une couleur existante, qui était la couleur de j boules a été choisie, et le nombre de boules de cette couleur devient $j + 1$. Dans ce cas, la probabilité de transition vaut alors

$$p(\bar{n}, n) = \frac{j\bar{n}_j}{\theta + k - 1}.$$

Le rapport des probabilités donné par la formule d'Ewens est par ailleurs :

$$\frac{P_\theta(n)}{P_\theta(\bar{n})} = \frac{k}{\theta + k - 1} \cdot \frac{j\bar{n}_j}{(j+1)n_{j+1}}.$$

Observons maintenant que

$$\begin{aligned} \sum_{\bar{n}} \frac{P(\mathbf{N} = \bar{n})}{P(\mathbf{N} = n)} p(\bar{n}, n) &= \frac{\theta}{\theta + k - 1} \cdot \frac{\theta + k - 1}{k} \cdot \frac{n_1}{\theta} \\ &+ \sum_{j=1}^{k-1} \frac{j\bar{n}_j}{\theta + k - 1} \cdot \frac{\theta + k - 1}{k} \cdot \frac{(j+1)n_{j+1}}{j\bar{n}_j}. \end{aligned}$$

En simplifiant les termes du membre de droite, nous obtenons finalement que

$$\sum_{\bar{n}} \frac{P(\mathbf{N} = \bar{n})}{P(\mathbf{N} = n)} p(\bar{n}, n) = \frac{n_1}{k} + \sum_{j=1}^{k-1} \frac{(j+1)n_{j+1}}{k} = 1,$$

puisque $\sum_{j=1}^k jn_j = k$. Ainsi donc, les probabilités définies par la formule d'Ewens satisfont (4.6.36). La distribution des couleurs de l'urne de Hoppe satisfait aussi cette propriété (4.6.36), avec la même valeur de la probabilité d'obtenir \bar{n} , par hypothèse de récurrence. On en déduit que $P(\mathbf{N} = n)$ est donnée par la formule d'Ewens. \square

On en déduit que

Proposition 4.6.11 *Si l'on considère la répartition de (N_1, \dots, N_h) , où h est un nombre fixé, alors que la taille k de l'échantillon tend vers l'infini, on a*

$$\lim_{k \rightarrow \infty} \mathbb{P}(N_1 = n_1, \dots, N_h = n_h) = \mathbb{P}(Y_1 = n_1, \dots, Y_h = n_h), \quad (4.6.37)$$

où les variables aléatoires Y_1, \dots, Y_h sont indépendantes et de loi une loi de Poisson de paramètres respectivement $\theta, \dots, \frac{\theta}{h}$.

4.6.7 Le point de vue processus de branchement avec immigration

On peut relier l'urne de Hoppe et la répartition allélique dans le coalescent à un processus de branchement linéaire avec immigration. Pour ce processus, les migrants arrivent dans la population aux temps successifs d'un processus de Poisson de paramètre θ . Chaque individu de la population suit les règles de reproduction d'un processus de branchement binaire. Les individus ne meurent jamais, et se reproduisent en donnant naissance à un autre individu après un temps exponentiel de paramètre 1. Si on regarde ce processus uniquement aux instants où le nombre de particules croît, on obtient un processus de branchement binaire en temps discret. Si on a une population de k individus, le nouvel arrivant aura un nouveau type avec probabilité $\frac{\theta}{\theta+k}$, et avec probabilité $\frac{k}{k+\theta}$, il prendra le type d'un individu choisi uniformément dans la population existante. À partir de cette description, il est clair que la chaîne de Markov incluse dans ce processus de branchement binaire avec immigration a même distribution que la répartition des couleurs dans le modèle de Hoppe.

Théorème 4.6.12 *Si chaque migrant est d'un nouveau type, et si les naissances sont du même type que celui des parents, alors la suite des états du processus de branchement avec immigration a la même distribution que celle générée par l'urne de Hoppe.*

Combinons cette dernière observation avec le fait suivant.

Lemme 4.6.13 (cf. Athreya-Ney [1] (p. 109)). *Soit un processus de branchement binaire (à taux 1) issu d'un individu unique. Alors le nombre de particules au temps t a une distribution géométrique avec probabilité de succès $p = e^{-t}$.*

Nous obtenons alors une nouvelle preuve du théorème 4.6.5.

Théorème 4.6.14 *Considérons le coalescent issu de l lignées et s'arrêtant quand il y en a k ($l \geq k$). Soit J_1, \dots, J_k les nombres de lignées dans l'échantillon de k individus, quand ils sont labellés au hasard. Alors (J_1, \dots, J_k) est uniformément distribué sur les vecteurs d'entiers positifs de somme l , et ainsi,*

$$\mathbb{P}(J_i = m) = \binom{l-m-1}{k-2} / \binom{l-1}{k-1}.$$

Preuve. Soit $Z_t^i, 1 \leq i \leq k$, des copies indépendantes du processus de branchement binaire. Si j_1, \dots, j_k sont des entiers positifs de somme l , alors par le lemme 4.6.13,

$$\mathbb{P}(Z_t^1 = j_1, \dots, Z_t^k = j_k) = (1-p)^{l-k} p^k \quad \text{où } p = e^{-t}.$$

Puisque le terme de droite dépend seulement de la somme l et du nombre de termes k , tous les vecteurs possibles ont la même probabilité. Comme déjà vu dans la preuve du

théorème 4.6.5, il y a $\binom{l-1}{k-1}$ vecteurs possibles (j_1, \dots, j_k) d'entiers positifs de somme l . Ainsi il s'en suit que

$$\mathbb{P} \left(Z_t^1 = j_1, \dots, Z_t^k = j_k \mid \sum_{j=1}^k Z_t^j = l \right) = 1 / \binom{l-1}{k-1}.$$

La distribution conditionnelle est donc uniforme sur tous les vecteurs possibles. Puisque le nombre de vecteurs (j_2, \dots, j_k) d'entiers positifs de somme $l - j_1$ est $\binom{l-j_1-1}{k-2}$, on obtient

$$\mathbb{P} \left(Z_t^1 = j_1 \mid \sum_{j=1}^k Z_t^j = l \right) = \binom{l-j_1-1}{k-2} / \binom{l-1}{k-1}.$$

Cela termine la preuve. □

Ce jeu de modèles entre le modèle en temps discret de l'urne de Hoppe et le modèle de branchement en temps continu avec immigration est très intéressant et utile. En particulier dans la preuve précédente, on a utilisé fortement l'indépendance des différentes familles créées par le processus de branchement. Cela va conduire à de jolis résultats sur le comportement asymptotique de la distribution d'Ewens, quand k est grand.

Définissons $S_j(k)$ la taille de la j -ème famille quand il y a k individus. (La j -ème plus ancienne). Alors on a le théorème suivant.

Théorème 4.6.15 *Pour $j \in \mathbb{N}^*$, on a $\lim_{k \rightarrow \infty} \frac{S_j(k)}{k} = X_j$, presque-sûrement, avec*

$$X_j = Z_j \prod_{i=1}^{j-1} (1 - Z_i),$$

où les variables aléatoires Z_j sont indépendantes et ont pour loi une loi beta($1, \theta$) de densité $\theta(1-x)^{\theta-1}$, $x \in [0, 1]$.

La loi de $(X_j)_j$ s'appelle la **loi de Poisson-Dirichlet**. C'est une loi sur les partitions de $[0, 1]$.

Bibliographie

- [1] K.B. Athreya, P.E. Ney. *Branching Processes*, Springer 1972.
- [2] M. Benaïm, N. El Karoui. *Promenade aléatoire. Chaînes de Markov et simulations ; martingales et stratégies. Cours de l'École Polytechnique.*
- [3] S. Méléard. *Aléatoire. Cours de l'École Polytechnique*
- [4] P. Billingsley : *Probability and Measure*, Wiley, New York (1979).
- [5] J.F. Delmas, B. Jourdain : *Modèles aléatoires : applications aux sciences de l'ingénieur et du vivant*, Springer 2006.
- [6] R. Durrett, *Probability Models for DNA Sequence Evolution*, Springer, 2nde édition, 2007.
- [7] N. El Karoui, E. Gobet : Modèles stochastiques en finance. Première partie : introduction au calcul stochastique. **Cours de l'École Polytechnique.**
- [8] W.J. Ewens : *Mathematical Population Genetics*. Second Edition. Springer, 2004.
- [9] W. Feller : *An introduction to Probability Theory and its Applications*, 2 Vol. Wiley, 1957.
- [10] P. Haccou, P. Jagers, V.A. Vatutin : *Branching Processes : Variation, growth and extinction of populations*, Cambridge University Press, 2005.
- [11] J. Istas : *Modèles mathématiques pour l'écologie. Cours de l'École Polytechnique.*
- [12] J. Istas : *Introduction aux modélisations mathématiques pour les sciences du vivant.* Springer 2000.
- [13] I. Karatzas, S.E. Shreve : *Brownian Motion and Stochastic Calculus*, Second Edition Springer, 1998.
- [14] S. Karlin, H.M. Taylor, *A Second Course in Stochastic Processes*, Academic Press 1981.
- [15] M. Kimmel, D.E. Axelrod : *Branching Processes in Biology*, Springer 2002.
- [16] M. Kot, *Elements of Mathematical Biology*, Cambridge 2001.
- [17] E. Pardoux, *Processus de Markov et Applications*, Dunod 2007.
- [18] B.L. Phillips, G.P. Brown, J.K. Webb, R. Shine : *Invasion and evolution of speed in toads.* Nature 439, 803, 2006.

- [19] E. Renshaw, *Modelling Biological Populations in Space and Time*, Cambridge University Press, 1991.
- [20] D. Serre, *Les matrices*, Dunod, 2001.