

الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي والبحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
جامعة عمار التليجي بالأغواط
UNIVERSITÉ AMAR TELIDJI LAGHOUAT
كلية العلوم
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES



Mémoire de MASTER

Domaine : Mathématiques et Informatique
Filière : Mathématiques
Option : Analyse Mathématique

Par : ELHANI Amina

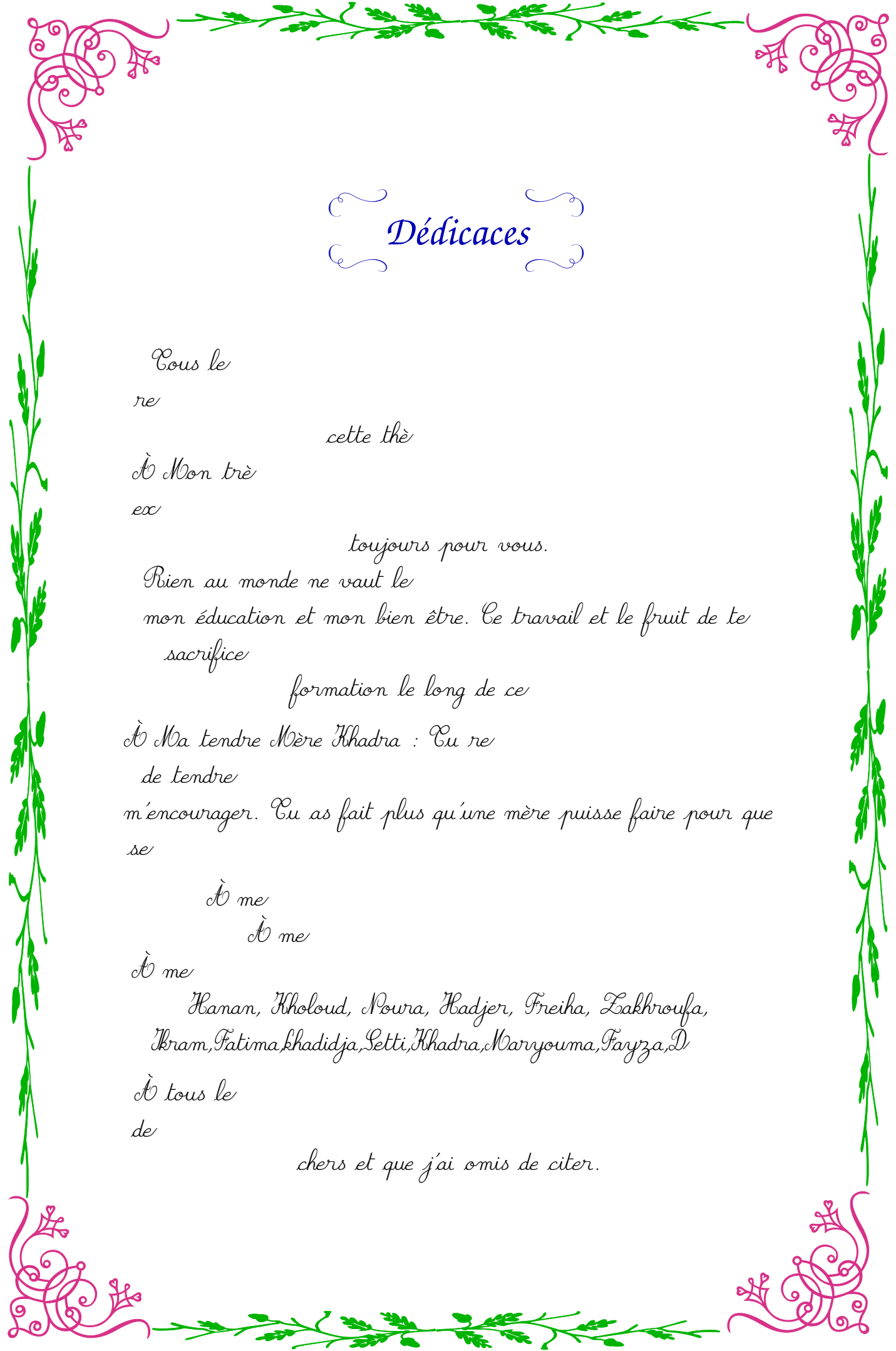
THÈME

Chaîne de markov et ses applications

Soutenu publiquement devant le jury composé de:

- | | |
|---------------------------------|-----------|
| • Mr. RAHMOUNE ABDELAZIZ. M.C.B | Président |
| • Mr. MERRAD BOULERBEH. M.A.A | Examineur |
| • Mr. BOUKEHILA AHCENE . M.C.B | Encadreur |

Année Universitaire 2020/2021



Dédicaces

Cous le
re

cette thè

À Mon trè
ex

toujours pour vous.

Rien au monde ne vaut le
mon éducation et mon bien être. Ce travail et le fruit de te
sacrifice

formation le long de ce

À Ma tendre Mère Khadra : Tu re
de tendre
m'encourager. Tu as fait plus qu'une mère puisse faire pour que
se

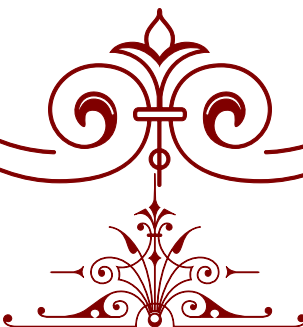
À me
À me

À me

Hanan, Kholoud, Noura, Hadjer, Freiha, Lakhroufa,
Ikram, Fatima Khadidja, Setti Khadra, Maryouma, Fayza, D

À tous le
de

chers et que j'ai omis de citer.



Remerciements



En tout premier lieu, je remercie mon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés et Je tiens à exprimer ma reconnaissance à Dr. Ahcene Boukehila pour avoir accepté de m'encadrer dans cette étude. Mes remerciements iront également aux membres du jury pour avoir accepté d'évaluer mon travail de recherche. A tous ceux qui m'ont aidée de près ou de loin , par un geste, une parole ou un conseil, je leur dis merci. Sans oublier tous nos enseignants qui nous ont assurés des études de haut niveau et qui nous permis d'acquérir des connaissances. Afin de n'oublier personne, mes vifs remerciements s'adressent à tous ceux qui m'ont aidée à la réalisation de ce modeste mémoire.



❖ *Amina*



Résumé

Le présent travail est un mémoire de master intitulée « Chaîne de Markov et ses applications » et composée de trois chapitres , le premier chapitre on étudie les différents propriétés de processus ainsi on donne la définition du processus de Markov et ses propriétés. Le deuxième chapitre est consacré à l'étude de la chaîne de Markov et la matrice de transition ; ses propriétés et les différents cas de cette chaîne de Markov, on étudie encore la chaîne de Markov dans un ensemble dénombrable et la chaîne de Markov dans le cas discret et continue . Le dernier chapitre représente l'application des chaînes de Markov représenté par le processus de Galton –Watson qui est une certaine classe de chaîne de Markov.

Mots-Clés : Chaîne de Markov ; matrice stochastique ; processus de Markov ; processus de Galton-Watson.

ملخص

هذه المذكرة بعنوان شبكة ماركوف وتطبيقاتها، وتتكون من ثلاثة فصول. نتناول في الفصل الأول الخصائص المختلفة للعمليات حتى نقدم عملية ماركوف وخصائصها. أما الفصل الثاني مخصص لدراسة شبكة ماركوف ومصفوفة الانتقال؛ أي الخصائص والحالات المختلفة لشبكة ماركوف، وندرسها في مجموعة معدودة وأيضاً في الحالة المنفصلة والمستمرة. يتمثل الفصل الأخير في تطبيق شبكة ماركوف حيث تمثلت في عملية جالتون-واتسون، وهي فئة معينة من شبكة ماركوف.

الكلمات المفتاحية: شبكة ماركوف وتطبيقاتها؛ عملية تطور ماركوف؛ مصفوفة عشوائية؛ عملية تطور جالتون-واتسون.

Abstract

The present work is a master memory entitled "Markov chain and its applications" and composed of three chapters, the first chapter we study the different properties of processes so we give the definition of the Markov process and its properties. The second chapter is devoted to the study of the Markov chain and the transition matrix; its properties and the different cases of this Markov chain, we still study the Markov chain in a countable set and the Markov chain in the discrete and continuous case. The last chapter represents the application of Markov chains represented by the Galton –Watson process which is a certain class of Markov chain.

Key words : Markov chain and its applications; Markov process; stochastic matrix; Galton_watson process.

Table des matières

Introduction Générale	5
1 Processus Aléatoire	8
1.1 Processus stochastique	8
1.2 Processus de Markov	13
2 Chaîne de Markov et matrice de transition	17
2.1 Introduction	17
2.2 Condition nécessaire et suffisante	19
2.3 Chaînes de Markov homogènes	23
2.4 Loi d'une chaîne de Markov	25
2.5 Chaînes de Markov sur un ensemble fini	28
2.6 Chaînes de Markov sur un ensemble dénombrable	35
2.7 Chaînes de Markov à temps discret	44
2.8 Chaînes de Markov à temps continu	45
2.9 Propriété de chaîne de Markov	47
3 Application de Chaînes de Markove	49
3.1 Le processus de Galton-Watson	49
Bibliographie	58



Introduction Générale

Une chaîne de Markov est un processus aléatoire $(X_n)_{n \in \mathbb{N}}$ dont les transitions sont données par une matrice stochastique $P(X_n, X_{n+1})$. Ces processus vérifient la propriété de Markov, c'est-à-dire qu'observé sa' partir d'un temps (d'arrêt) T , $(X_{T+n})_{n \in \mathbb{N}}$ ne dépend que de X_T et est de nouveau une chaîne de Markov. Les états d'une chaîne de Markov peuvent être classés en deux catégories : les états transitoires, qui ne sont visités qu'un nombre fini de fois presque (*p.s.*), et les états récurrents, qui une fois atteints sont visités *p.s.* une infinité de fois, ainsi que tous les autres états dans la même classe de récurrence. Pour une chaîne de Markov irréductible récurrente et la loi marginale du processus convergent soit vers l'unique mesure de probabilité P -invariante, soit vers le vecteur nul (récurrence nulle). Cette théorie s'applique en particulier aux marches aléatoires et aux modèles de files d'attente [7].

Un processus stochastique est un phénomène qui évolue dans le temps d'une manière aléatoire. La vie quotidienne et la science nous donnent beaucoup d'exemples de ce genre de phénomène, ou' en tout cas des phénomènes qui peuvent être compris de cette façon. La météo ; la population d'une ville ; le nombre de personnes dans une file d'attente et la position d'une particule de pollen dans un fluide, sont des exemples de processus stochastiques. Ce dernier fut étudié pour la première fois par le botaniste Robert Brown en 1827 et reçu le nom de mouvement brownien. Il joue un rôle fondamental dans la théorie des processus aléatoires, un peu comme la distribution gaussienne dans la théorie des probabilités. En 1920, N. Wiener donne une définition mathématique du mouvement brownien. Il étudie en particulier ses trajectoires qui sont continues mais nulle part différentiables (à un moment la vitesse ne peut être définie car les changements de direction sont trop rapides). Dans les années quarantes, K. Itô développa un calcul différentiel spécifique au mouvement brownien [5].



La fiabilité est née, sans doute, au cours de la 2^{ème} guerre mondiale. Actuellement, l'étude de fiabilité devient une discipline à part entière. Les problèmes économiques (coût de défaillance, gestion de personnel de maintenance, etc) rendent nécessaire la connaissance de fiabilité du système. L'analyse de la fiabilité constitue une phase indispensable dans toute étude de sûreté de fonctionnement d'un système. Il est à signaler que l'analyse de fiabilité dans n'importe quel domaine est un outil très important pour caractériser le comportement d'un système dans les différentes phases de vie.

La théorie de la fiabilité a pour objectif d'étudier l'aptitude de dispositifs techniques (machines, équipements,...), à accomplir une fonction requise, dans des conditions données, durant un temps donné. Actuellement, c'est une discipline à part entière. Prévoir la fiabilité d'un système est essentielle pour des problèmes de sécurité : systèmes de freinage, systèmes nucléaires, systèmes informatiques...). La quasi-impossibilité de réparer certains matériels (satellites), les problèmes économiques : coût des défaillances, gestion du personnel de maintenance, maintenance des stocks des pièces de rechange...) rendent nécessaire la connaissance de la fiabilité des systèmes utilisés.

Les défaillances se produisant généralement de façon aléatoire, il est logique de faire appel au calcul des probabilités pour étudier des problèmes de fiabilité. Ainsi, nous définissons la fiabilité d'un dispositif comme étant sa probabilité de fonctionner correctement pendant une durée donnée, ou, ce qui revient au même, la probabilité qu'aucune défaillance ne se produise pendant cette durée ([1]).

Processus Aléatoire

Introduction

Les processus aléatoires décrivent l'évolution d'une grandeur aléatoire en fonction du temps. Il existe de nombreuses applications des processus aléatoires notamment en physique statistique (par exemple le ferromagnétisme, les transitions de phases, etc.), en biologie (évolution, génétique et génétique des populations), médecine (croissance de tumeurs, épidémie), et bien entendu les sciences de l'ingénieur. Dans ce dernier domaine, les applications principales sont pour l'administration des réseaux, de l'Internet, des télécommunications et bien entendu dans les domaines économique et finance [6].

1.1 Processus stochastique

Définition 1.1

On peut définir un processus stochastique comme étant une famille $(X_t)_{t \in T}$ de variables aléatoires indexées par le temps t . Les mots processus et stochastique signifient respectivement fonction et aléatoire.

Alors qu'une variable aléatoire X associe à chaque $\omega \in \Omega$ une réalisation $X(\omega)$ un processus stochastique $(X_t)_{t \in T}$ associe à chaque une fonction (ou trajectoire) $(X_t(\omega))_{t \in T}$; $T \rightarrow E$; $t \rightarrow X_t(\omega)$ où E est l'espace d'arrivée des variables aléatoires $\{X_t\}$. Passer des variables aléatoires aux processus stochastiques revient à passer en analyse des points aux fonctions. itre d'exemple, la trajectoire d'une mouche en fonction du temps peut être modélisée par un processus stochastique aleurs dans $E = \mathbb{R}^3$ Lorsque l'ensemble des temps T est au plus dénombrable (par exemple $T = \mathbb{N}$), on parle de processus stochastiques á temps discret. Lorsqu'il est continu (i.e. $T = [0, t_0]$ ou $T = \mathbb{R}^+$), on parle de processus stochastiques á temps continu Dans ce travail, on abrège les expressions "variable aléatoire" en v.a. et "indépendantes et identiquement distribuées en i.i.d. Les situations réelles pouvant être modélisées par des processus stochastiques sont nombreuses ([1]).

Définition 1.2

On appelle processus stochastique une famille $X_t, t \in T$ de variables aléatoires définies dans le même espace de probabilité $(\Omega; F; P)$ et à valeurs dans l'espace mesurable $(E; \xi)$, $t \in T$ représente une date. Lorsque $T \subseteq \mathbb{Z}$, on parlera de processus à temps discret (suite stochastique) on notée $(X_n, n \in \mathbb{N})$ lorsque T est un intervalle $I \subseteq \mathbb{R}$, on parlera de processus à temps continu [5].

Définition 1.3

Un processus (aléatoire ou stochastique) rend compte de l'évolution, d'un phénomène aléatoire; au cours du temps,. La réalisation d'un processus est appelée trajectoire. On note $X(t)$ l'état du phénomène au temps t . $X(t)$ est une variable aléatoire.

- La loi de $X(t)$ dépend en général de t .
- Pour deux dates t_1 et t_2 quelconques, $X(t_1)$ et $X(t_2)$ ne sont, en général, pas indépendantes.

Un processus prend ses valeurs dans un espace des états et évolue dans un espace des temps ([10]).

Espace des temps

On note T l'ensemble dans lequel évolue le temps. On peut distinguer :

- les processus à temps discret pour lesquels les temps sont des entiers successifs : $T = 0, 1, \dots, m, \dots$. On note X_m la valeur du processus à la date $t = m$
- les processus à temps continu pour lequel T est inclus dans $\mathbb{R}$. On note X_t la valeur du processus à la date t . Cet espace T est parfois appelé espace des indices. On peut en effet définir des processus évoluant non pas dans le temps mais dans l'espace; dans ce cas T peut avoir plusieurs dimensions (exemple : hauteur des vagues en fonction de la latitude et de la longitude).

On n'étudiera ici que des processus temporels pour lesquels T a une seule dimension ([10]). Espace des états.

On note E l'espace dans lequel les variables $X(t)$ prennent leurs valeurs. On peut distinguer : les espaces des états dénombrables, finis ou infinis; les espaces des états continus. Le mouvement brownien qui décrit la trajectoire d'une particule au cours du temps est un exemple de processus à espace des états continus.

On n'étudiera ici que des processus pour lesquels E est dénombrable, et le plus souvent fini ([10]).

Exemples de processus aléatoires

Exemple 1.1

a) **Signal télégraphique :**

Ce processus, utilisé en théorie de la communication, rend compte de l'état d'occupation d'une ligne

A_k : instant du début de la K^{ime} communication ;

D_k : instant du la fin de la K^{ime} communication ;

$$T = \mathbb{R}_+; E = \{-1, 1\}$$

$X_t = 1$ si la ligne est libre, c'est-à-dire s'il existe k tel que $D_k \leq t < A_{k+1}$

$X_t = -1$ si la ligne est occupée, c'est-à-dire s'il existe k tel que $A_k \leq t < D_k$ On

a là, un processus (dit à créneaux), markovien, à accroissements indépendants, homogène dans le temps.

b) **Processus de ramification** Ce processus est utilisé pour suivre l'évolution de certaines populations animales ou cellulaires, ou bien d'un nom chez les humains.

Chaque individu génère, indépendamment des autres, un nombre aléatoire de descendants. X_n est le nombre d'individus total à la n^{ime} génération

$$E = \mathbb{N}; T = \mathbb{N}$$

Le processus est markovien.

c) **Files d'attente :** Ces processus sont utilisés en recherche opérationnelle. Des clients se présentent à des guichets à des temps aléatoires, pour y recevoir des services de durée aléatoire (ex : banque, magasin, péage d'autoroute...) X_t est le nombre de clients en attente à l'instant t .

$$E = \mathbb{N}; T = \mathbb{R}$$

d) **Mouvement Brownien :** Dans un gaz, chaque particule est soumise aux impacts incessants de ses voisines. Les collisions provoquent des déplacements. X_t est la position d'une particule donnée à l'instant t . $X_t - X_s$ est le déplacement pendant $[t; s]$: c'est la somme d'un grand nombre de petits déplacements et il sera légitime de supposer que sa loi est une loi Normale.

$$E \subset \mathbb{R} \quad ; \quad T = \mathbb{R}_+.$$

Le processus est à accroissements indépendants, homogène dans le temps. Ce travail a pour objectif de donner quelques éléments de théorie sur les processus

aléatoires généraux : processus de Markov, de Poisson, de naissance et de mort, puis sur les systèmes d'attentes : files uniques et réseaux, avec entre temps une application à la fiabilité ([8]).

Exemple 1.2

1. Auto fécondations :

ε est l'ensemble (fini) des génotypes possibles et $\mathbb{T}$ est discret puisqu'il représente les générations successives. Exemple : $E = \text{homozygote, hétérozygote}$, $\mathbb{T} = 0, 1, 2, \dots, X_m$ est le génotype du même descendant (chaque génération est composée d'un seul individu, voir Fig.1.1,

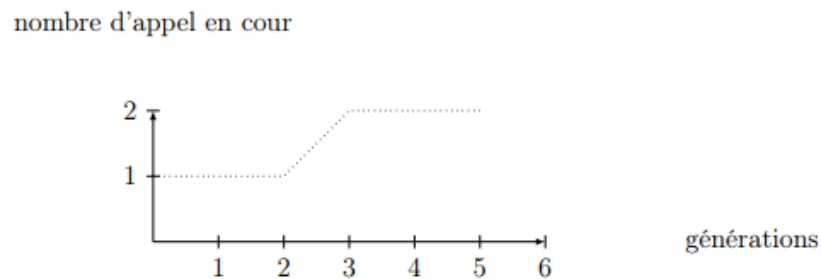


FIGURE 1.1 – Génotypes d'individus obtenus par auto fécondations successives

2. Appels téléphoniques :

ε est fini car la capacité du standard est finie et $\mathbb{T}$ est continu car les appels peuvent survenir à n'importe quel moment. Exemple : $\varepsilon = 0, 1, \dots, 10$, $\mathbb{T} = [8h; 17h30]$, $X(t)$ est le nombre d'appels en attente au temps t (voir Fig.1.2).

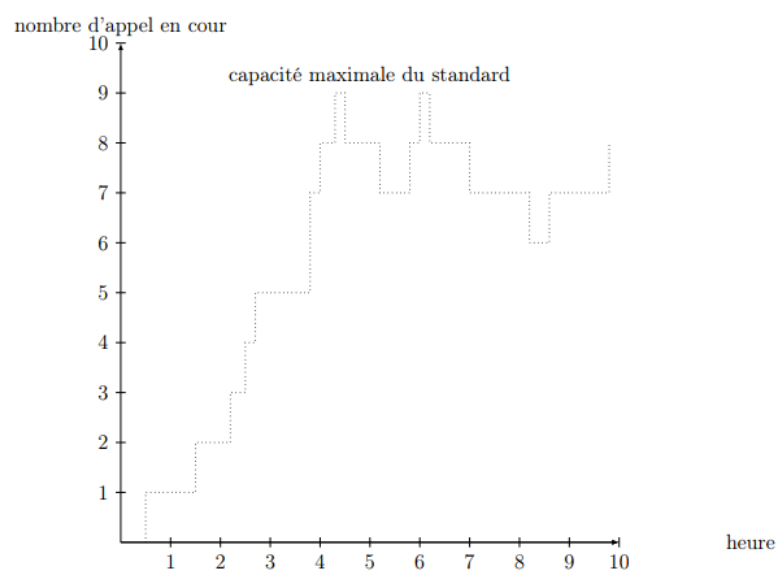


FIGURE 1.2 – Nombre d'appels en attente à un standard téléphonique

3. Taille d'une famille :

ε est infini (dénombrable) car le nombre de descendants n'est pas limité et $\mathbb{T}$ est discret puisqu'il représente les générations successives. Ici $\varepsilon = \mathbb{T} = \mathbb{N}$, X_m est le nombre de descendants au bout de m générations (voir Fig.1.3)

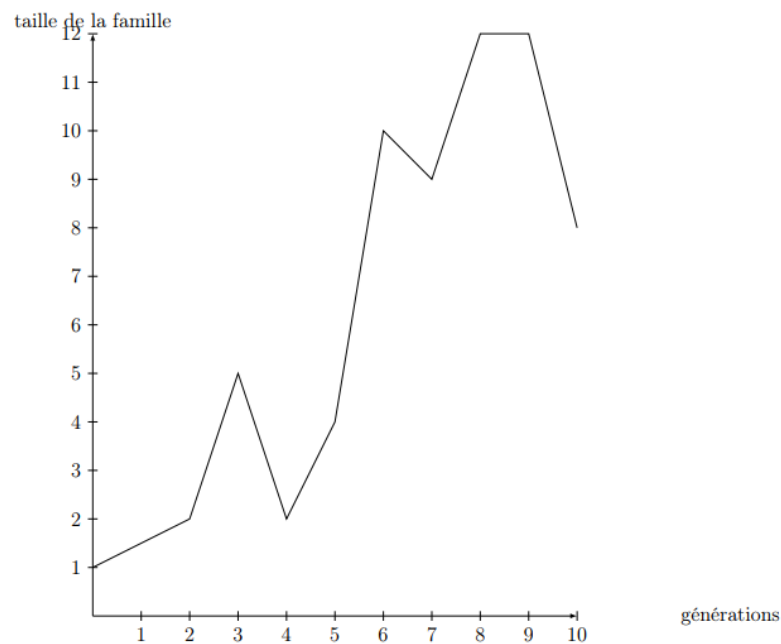


FIGURE 1.3 – Taille d'une famille au cours des générations successives

4. Dynamique d'une population :

ε est infini si on suppose que la taille de la population n'est pas limitée et $\mathbb{T}$ est continu car les individus peuvent naître ou mourir à n'importe quel moment. Ici $\varepsilon = \mathbb{N}$ et $\mathbb{T} = [0; +\infty[$, $X(t)$ est la taille de la population au temps t [10].

1.2 Processus de Markov

Définition 1.4

Le processus $(X_t)_{t \geq 0}$ est dit de Markov, si

a) **axiome de Markov :**

pour tout $t_1 < t_2 < \dots < t_n < t_{n+1}$. pour tout $x_1, x_2, \dots, x_{n+1} \in E$:

$$P([X_{t_{n+1}} = x_{n+1}] / [X_{t_1} = x_1] \dots \cap [X_{t_n} = x_n]) = P([X_{t_{n+1}} = x_{n+1}] / [X_{t_n} = x_n])$$

b) **axiome d'homogénéité :**

pour tous s et t , pour tous $x, y \in E$, $P([X_{t+s} = y] / X_s = x)$ ne dépend que de t (et non des instants s et $t + s$).

Propriétés :

1. $P(t)$ est une matrice stochastique, .i.e. $p_{x,y}(t) \geq 0$ et $\sum_{y \in E} p_{x,y}(t) = 1$ pour tout x .
2. Pour tout s et pour tout t , $P(s+t) = P(s)P(t)$
3. Pour tout s et pour tout t , $\pi(s+t) = \pi(s)P(t)$

Démonstration :

1. $P_{x,y}(t) \in [0, 1]$ car c'est une probabilité. De plus, la ligne x correspond à la loi $P_{X_t}^{[X_0=x]}$ et on a bien

$$\sum_{y \in E} p_{x,y}(t) = \sum_y P^{[X_0=x]}([X_t = y]) = 1.$$

2. Pour calculer $P_{x,y}(t+s) = P^{X_0=x}([X_{t+s} = y])$ on va faire intervenir les différents états pouvant être occupés à l'instant t :

$$\begin{aligned} P^{X_0=x}([X_{t+s} = y]) &= \frac{P([X_0 = x] \cap [X_{t+s} = y])}{P([X_0 = x])} \\ &= \sum_{z \in E} \frac{P([X_0 = x] \cap [X_t = z] \cap [X_{t+s} = y])}{P([X_0 = x])} \\ &= \sum_{z \in E} \frac{P([X_{t+s} = y] / [X_0 = x] \cap [X_t = z]) P([X_0 = x] \cap [X_t = z])}{P([X_0 = x])} \\ &= \sum_{z \in E} P([X_{t+s} = y] / [X_t = z]) P([X_t = z]) / [X_0 = x] \end{aligned}$$

d'après l'axiome de Markov (X_0 n'apporte rien de plus que X_t pour déterminer la loi de X_{t+s}).

Ainsi, on a

$$P_{x,y}(t+s) = \sum_{z \in E} P_{x,z}(t) P_{z,y}(s),$$

qui est le coefficient (x, y) de la matrice produit $P(t)P(s)$.

$$3. \pi_y(t+s) = P([X_t + s = y]) = \sum_{x \in E} P([X_t + s = y] / [X_t = x]) P([X_t = x]) \text{ c'est-à-dire } \pi_y(t+s) = \sum_{x \in E} \pi_x(t) p_{x,y}(s) \text{ [8].} \quad \blacksquare$$

Types de processus de Markov

On peut également distinguer deux types d'espace états, selon qu'il est discret ou continu comme respectivement dans les types 1 et 3 et les types 2 et 4. [1] On appelle espace des états (des phases) l'ensemble E où les variables X_n prennent leurs valeurs. L'ensemble E peut être discret ou continu. Par conséquent, on distingue quatre types de processus :

- processus Un espace des temps T discret
- processus Un espace des temps T continue
- processus Un espace des états E discret
- Processus Un espace des états E non discret [5]

1. Processus un Espace d'états E discret :

Lorsque les variables aléatoires successives sont des variables discrètes munies d'une fonction de probabilité, on parle de chaîne de Markov.

Bien que les chaînes de Markov s'appliquent à des phénomènes dont l'aspect temporel est généralement sans intérêt, on peut associer aux valeurs successives X_i , les instants t_i . La propriété markovienne selon laquelle la probabilité d'un état du système ne dépend que de son état précédent à travers une probabilité conditionnelle appelée probabilité de transition s'exprime par :

$$\mathbb{P}(x_n, t_n \mid x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \dots; x_1, t_1; x_0, t_0) = \mathbb{P}(x_n, t_n \mid x_{n-1}, t_{n-1})$$

$$t_n > t_{n-1} > \dots > t_1 > t_0$$

Une chaîne de Markov est entièrement définie par la probabilité au premier ordre $P(x, t)$ et la probabilité de transition. On obtient par exemple la probabilité au second ordre par :

$$\mathbb{P}(x_1, t_1; x_2, t_2) = \mathbb{P}(x_1, t_1) P(x_2, t_2 \mid x_1, t_1) t_2 > t_1.$$

Elle est donc également définie par la probabilité au second ordre. Enfin, elle peut être définie par l'état initial et la probabilité de transition.

C'est-à-dire fini ou dénombrable. Il sera, dans ce cas, souvent pratique d'identifier E avec une partie de $\mathbb{N}$ ou de $\mathbb{Z}$ [5].

2. processus un espace d'états E continu :

Les chaînes de Markov trouvent des applications dans les domaines les plus divers mais les processus considérés dans les problèmes dynamiques, en particulier en vibrations, portent généralement sur des variables aléatoires continues.

Dans ces conditions, la probabilité d'obtenir une valeur donnée est généralement nulle et les probabilités d'apparition doivent être remplacées par des densités de probabilité dans la formule de la propriété markovienne :

$$P(x_n, t_n \mid x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \dots; x_1, t_1; x_0, t_0) = P(x_n, t_n \mid x_{n-1}, t_{n-1}) \\ t_n > t_{n-1} > \dots > t_1 > t_0$$

par exemple $E = \mathbb{R}$ ou $E \subset \mathbb{R}^2$ (partie du plan) ou $E \subset \mathbb{R}^3$ (partie de l'espace)

3. Processus un espace des temps T continu :

$T = \mathbb{R}_+$: le processus est dit continu on garde les yeux fixés sur un système qui évolue dans le temps à partir d'un instant t_0 que l'on prend pour origine des temps $t = 0$.

4. Processus un espace des temps T discret :

$T = \mathbb{N}$: le processus est dit discret ; on regarde ce qu'il se passe à chaque unité de temps, ou bien on fait une suite d'opérations et on regarde ce qu'il se passe à chaque opération (ex : lancer d'une pièce). [5].

Exemple de Processus de Markov

1. La transmission du patrimoine génétique est typiquement un phénomène markovien puisque toute l'information apportée par les ancêtres est résumée dans le génotype des parents. C'est un phénomène homogène dans le temps puisque les lois qui gouvernent cette transmission ne varient pas au cours des générations.
2. Si les appels surviennent indépendamment les uns des autres, le nombre d'appels reçus durant un petit intervalle ne dépend pas du nombre d'appel déjà enregistré ; dans ce cas, le processus peut être considéré comme markovien . Par contre, il n'est pas homogène car la fréquence des appels peut varier au cours de la journée (heures de pointe, pause du midi, etc.).
3. Si la loi du nombre de descendants d'un individu est la même à toutes les générations, alors l'évolution de la taille d'une famille est un processus markovien homogène. Cet exemple sera traité en détail.
4. Le processus décrivant l'évolution de la taille d'une population est appelé "processus de naissances et morts". Pour la même raison que celle invoquée, on peut considérer qu'il est markovien si les naissances et les morts ont lieu indépendamment les uns

des autres. Il n'est, en général, pas homogène au cours du temps, par exemple si on considère l'influence du cycle des saisons sur les reproductions.

5. Le cours d'une action est fréquemment présenté comme un exemple de processus ayant une mémoire longue. Les évolutions ne dépendent pas seulement de la dernière valeur du titre, mais aussi des évolutions passées (phases de croissance, chocs boursiers, etc.) Dans ce cas, l'hypothèse de Markov n'est pas raisonnable. L'homogénéité temporelle peut elle aussi être mise en doute selon le modèle économique sous-jacent. on regarde ce qu'il se passe à chaque opération (ex : lancer d'une pièce). ([10]).

Chaîne de Markov et matrice de transition

2.1 Introduction

Une chaîne de Markov est une suite de variables aléatoires $(X_n, n \in N)$ qui permet de modéliser l'évolution dynamique d'un système aléatoire : X_n représente l'état du système à l'instant n . La propriété fondamentale des chaînes de Markov, dite propriété de Markov, est que son évolution future ne dépend du passé qu'au travers de sa valeur actuelle. Autrement dit, conditionnellement à $X_n, (X_0, \dots, X_n)$ et $(X_{n+k}, k \in N)$ sont indépendants. Les applications des chaînes de Markov sont très nombreuses (réseaux, génétique des populations, mathématiques financières, gestion de stock, algorithmes stochastiques d'optimisation, simulation, ...) [3].

Définition 2.1

on appelle matrice de transition, la matrice $P = (p_{i,j})_{i,j \in E}$

$$P = \begin{pmatrix} p_{i_0,j_0} & p_{i_0,j_1} & p_{i_0,j_2} & \cdots & \cdots \\ p_{i_1,j_0} & p_{i_1,j_1} & p_{i_1,j_2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

la matrice de transition P de la chaîne discret (X_n) est une matrice carrée, constituée par les probabilités de transition . elle vérifie les propriétés : pour tout couple (i, j) de E :

$$0 \leq p_{i,j} \leq 1$$

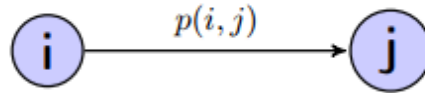
pour tout $i \in E$ on a

$$\sum_{j \in E} p_{i,j} = 1$$

[7]

Lien avec les graphes

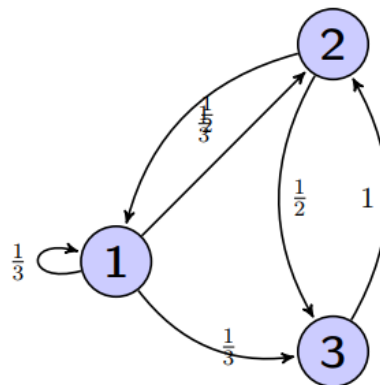
La matrice de transition d'une chaîne de Markov homogène finie peut être associée à un graphe dont les sommets sont les états. Deux états i et j sont reliés par un arc



de valeur $p(i, j)$ si $p(i, j)$ est non nul. Par exemple, la chaîne associée à la matrice P :

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

a le graphe associé présenté en Fig



2.2 Condition nécessaire et suffisante

Pour simplifier les notations on notera $X_{0:n}$ au lieu de $(X_0; \dots; X_n)$ et $x_{0:n}$ au lieu de $(x_0; \dots; x_n)$

Proposition 2.1

Soit $\{X_n; n \geq 0\}$ une suite de variables aléatoires. $\{X_n; n \geq 0\}$ est une chaîne de Markov de matrice de transition P et de loi initiale μ si et seulement si pour tout $n \in N$ la loi du vecteur aléatoire $(X_0; \dots; X_n)$ est donnée par :

$$P(X_0 = x_0; \dots; X_n = x_n) = \mu(x_0) P(x_0; x_1) \dots P(x_{n-1}; x_n) :$$

Démonstration :

$\implies$ Supposons que $\{X_n; n \geq 0\}$ soit une chaîne de Markov de matrice de transition P et de loi initiale μ . Soit $n \in N$. On a

$$P(X_{0:n} = x_{0:n}) = P(X_n = x_n \mid X_{0:n-1} = x_{0:n-1}) P(X_{0:n-1} = x_{0:n-1}) :$$

Puis par définition d'une chaîne de Markov,

$$P(X_{0:n} = x_{0:n}) = P(X_n = x_n \mid X_{n-1} = x_{n-1}) P(X_{0:n-1} = x_{0:n-1})$$

Comme P est la matrice de transition de $\{X_n; n \geq 0\}$

$$P(X_{0:n} = x_{0:n}) = P(x_{n-1}; x_n) P(X_{0:n-1} = x_{0:n-1}).$$

On obtient par itérations :

$$P(X_{0:n} = x_{0:n}) = P(x_{n-1}; x_n) \dots P(x_0; x_1) P(X_0 = x_0).$$

Enfin, on conclut en utilisant le fait que μ est la loi initiale de $\{X_n; n \geq 0\}$:

$$P(X_{0:n} = x_{0:n}) = P(x_{n-1}; x_n) \dots P(x_0; x_1) \mu(x_0).$$

$\Leftarrow$ Supposons maintenant que pour tout $n \in N$ la loi de $(X_0; \dots; X_n)$ est donnée par :

$$P(X_0 = x_0; \dots; X_n = x_n) = \mu(x_0) P(x_0; x_1) \dots P(x_{n-1}; x_n).$$

Soient $k \in N$ et $(x_0; \dots; x_{k-1}; x; y) \in E^{k+2}$.

$$\begin{aligned} P(X_{k+1} = y \mid X_{0:k} = (x_0; \dots; x_{k-1}; x)) &= \frac{P(X_{0:k+1} = (x_0; \dots; x_{k-1}; x; y))}{P(X_{0:k} = (x_0; \dots; x_{k-1}; x))} \\ &= \frac{(x_0) P(x_0; x_1) \dots P(x_{k-1}; x) P(x; y)}{\mu(x_0) P(x_0; x_1) \dots P(x_{k-1}; x)} \\ &= P(x; y). \end{aligned}$$

Donc $\{X_n; n \geq 0\}$ est une chaîne de Markov homogène de matrice de transition P . L'hypothèse nous donne aussi avec $n = 0$ et $x \in E$, $P(X_0 = x) = \mu(x)$. Donc $\{X_n; n \geq 0\}$ est une chaîne de Markov de matrice de transition P et de loi initiale μ . ■

Proposition 2.2

Soit $\{X_n; n \geq 0\}$ une chaîne de Markov de matrice de transition P . La matrice P est dite stochastique, c'est à dire qu'elle vérifie les deux propriétés suivantes :

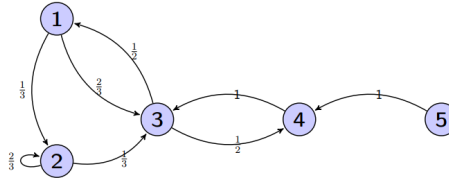
- (a) $\forall x; y \in E, P(x; y) \geq 0$
- (b) $\forall x \in E, \sum_{y \in E} P(x; y) = 1$: [3]

Exemple 2.1

la matrice de transition est :

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

le graphe associe :



[5].

Définition 2.2

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. à valeurs dans un espace E fini ou dénombrable, appelé espace d'états. On dit que $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov si :

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

pour tout entier $n \in \mathbb{N}$, pour tout état j et pour toute suite d'états $i_0, i_1, \dots, i_{n-1}$, i pour lesquels la probabilité conditionnelle a un sens, i.e.

$$P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) > 0$$

Si de plus la quantité $P(X_{n+1} = j \mid X_n = i)$ ne dépend pas de n , i.e.

$$P(X_{n+1} = j \mid X - n = i) = P(X - 1 = j \mid X - 0 = i)$$

alors la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ est dite homogène.

Il faut comprendre une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ comme une promenade dans l'espace d'états E , la variable X_n indiquant l'état dans lequel on est à l'instant n . La v.a. X_0 représente l'état initial duquel démarre la chaîne. Selon le contexte, X_0 pourra être aléatoire ou déterministe. La propriété de Markov signifie que, connaissant le dernier état visité (disons à l'instant n), la loi du prochain état visité (i.e. la loi de X_{n+1}) ne dépend pas des états visités depuis l'instant 0 jusqu'à l'instant $n - 1$. Plus prosaïquement, on dit que conditionnellement au présent, le futur ne dépend pas du passé.

Mais il dépend du présent : X_n et X_{n+1} n'ont aucune raison d'être indépendantes ! La propriété d'homogénéité d'une chaîne de Markov exprime quant à elle que la probabilité d'aller de i en j reste la même au cours du temps. Elle permet de regrouper en une seule matrice (indépendante de n) les probabilités de transition entre deux états quelconques.

Définition 2.3

Une matrice $P = (P(x, y), x, y \in E)$ est dite matrice stochastique si ses coefficients sont positifs et la somme sur une ligne des coefficients est égale à 1 :

$$P(x, y) \geq 0 \text{ et } \sum_{z \in E} P(x, z) = 1, \text{ pour tous } x, y \in E$$

Définition 2.4

Soit P une matrice stochastique sur E . Une suite de variables aléatoires $(X_n, n \in \mathbb{N})$ à valeurs dans E est appelée chaîne de Markov de matrice de transition P si pour tous $n \in \mathbb{N}, x \in E$, on a

$$P(X_{n+1} = x \mid X_n, \dots, X_0) = P(X_{n+1} = x \mid X_n) = P(X_n, x)$$

On dit que la chaîne de Markov est issue de μ_0 si la loi de X_0 est μ_0 . Comme l'espace d'état est discret l'équation est équivalente à : pour tous $x_0, \dots, x_n \in E$, tels que $P(X_n = x_n, \dots, X_0 = x_0) > 0$,

$$P(X_{n+1} = x \mid X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = x \mid X_n = x_n) = P(x_n, x)$$

Si $P(X_0 = x) = 1$, autrement dit μ_0 est la masse de Dirac en x , on dira plus simplement que la chaîne de Markov est issue de x .

Exemple

Transmission d'un bit informatique

Un bit informatique valant 0 ou 1 est transmis d'un poste A vers un poste B en passant par N intermédiaires. Chacun de ces intermédiaires transmet correctement le bit

avec probabilité $\mathbb{P}$ et l'inverse avec probabilité $1 - \mathbb{P}$, indépendamment les uns des autres. Le bit (aléatoire) d'entrée, disons X_0 , est supposé indépendant des intermédiaires. Pour $n = 1, \dots, N$, notons X_n le bit intermédiaire. La suite de v.a. $(X_n)_{0 \leq n \leq N}$ est à valeurs dans l'espace d'états $\mathbb{E} = 0, 1$. Vérifions que c'est une chaîne de Markov. Considérons pour ce faire, une suite $i_0, \dots, i_n, i_{n+1}$ d'éléments de $0, 1$

$$\begin{aligned} \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{\mathbb{P}(X_{n+1} = i_{n+1} \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \frac{\mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n| \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n|) \end{aligned}$$

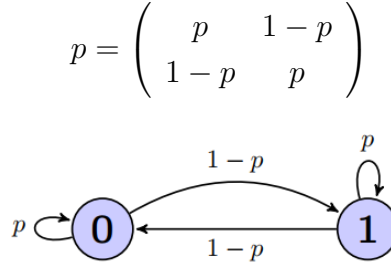
du fait de l'indépendance entre la v.a. $X_{n+1} - X_n$ (qui représente l'action du $(n+1)^e$ intermédiaire) et l'état du bit à la sortie des n premiers intermédiaires. Pour la même raison ;

$$\begin{aligned} \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{\mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n|) \mathbb{P}(X_n = i_n)}{\mathbb{P}(X_n = i_n)} \\ &= \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n) \end{aligned}$$

Enfin, le caractère homogène de la chaîne $(X_n)_{0 \leq n \leq N}$ résulte du calcul :

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(|X_{n+1} - X_n| = |j - i|) = \begin{cases} p & : \text{si } i = j \\ 1 - p & : \text{sinon} \end{cases}$$

Voici la matrice et le graphe de transition de cette chaîne :



La marche aléatoire simple

Soient $(Y_n)_{n \geq 1}$ une suite de v.a. i.i.d., chacune valant 1 avec probabilité p et -1 avec probabilité $1 - p$. Soit $Y_0 \in \mathbb{Z}$ une v.a. indépendante des Y_n , représentant le point de départ sur l'axe Z de la chaîne. On définit la marche aléatoire simple par

$$\begin{aligned} X_{n+1} &= X_n + Y_{n+1} \\ &= Y_0 + Y_1 + \dots + Y_n + Y_{n+1} \end{aligned}$$

pour tout $n \in \mathbb{N}$ Le processus stochastique $(X_n)_{n \geq 1}$ est une chaîne de Markov homogène. Son espace d'états $\mathbb{E} = \mathbb{Z}$ est cette fois infini. Comme dans l'exemple du bit informatique, le caractère markovien provient de l'indépendance des Y_n :

$$\begin{aligned}
P(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{P(X_{n+1} = i_{n+1} \text{ et } X_n = i_n, \dots, X_0 = i_0)}{P(X_n = i_n, \dots, X_0 = i_0)} \\
&= \frac{P(Y_{n+1} = i_n + 1 - i_n \text{ et } X_n = i_n, \dots, X_0 = i_0)}{P(X_n = i_n, \dots, X_0 = i_0)} \\
&= P(Y_{n+1} = i_{n+1} - i_n) \\
&= \frac{P(Y_{n+1} = i_n + 1 - i_n) P(X_n = i_n)}{P(X_n = i_n)} \\
&= \frac{P(X_{n+1} - X_n = i_{n+1} - i_n \mid X_n = i_n)}{P(X_n = i_n)} \\
&= P(X_{n+1} = i_{n+1} \mid X_n = i_n) [1]
\end{aligned}$$

2.3 Chaines de Markov homogènes

Définition 2.5

Soit $\{X_n, n \geq 0\}$ une suite de variables aléatoires sur $(\Omega; F; P)$ à valeurs dans E . $\{X_n, n \geq 0\}$ est une chaîne de Markov si pour tout $n \in N$ et pour tous $(x_0; \dots; x_{n+1}) \in \mathbb{E}^{n+1}$ on a $\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n; \dots; X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)$ E est appelé l'espace des états. Pour $n \in \mathbb{N}$, X_n est l'état de la chaîne à l'instant n . En particulier, X_0 est l'état initial de la chaîne.

Remarque 2.1

Une réalisation du vecteur aléatoire $(X_0; \dots; X_n)$ est parfois appelée trajectoire de la chaîne de Markov $\{X_n, n \geq 0\}$.

Remarque 2.2

La propriété ci-dessus s'appelle la propriété de Markov. On peut aussi l'énoncer de cette façon : conditionnellement au présent, le futur et le passé sont indépendants.

Définition 2.6

Soit $\{X_n, n \geq 0\}$ une chaîne de Markov d'espace d'états E . $\{X_n, n \geq 0\}$ est dite homogène si

$$P(X_{n+1} = y \mid X_n = x)$$

ne dépend que de x et de y . Dans toute la suite, on ne considérera que des chaînes de Markov homogènes d'espace d'états E .

Définition 2.7

Soit $\{X_n, n \geq 0\}$; une chaîne de Markov telle que pour tous $x, y \in \mathbb{E}$,

$$\mathbb{P}(X_1 = y \mid X_0 = x) = P(x; y) \text{ et } \mathbb{P}(X_0 = x) = \mu(x) . \text{ Alors :}$$

- (a) la matrice P définie par $P = (P(x; y))_{x, y \in \mathbb{E}}$ est la matrice de transition de la chaîne de Markov.
- (b) le vecteur μ défini par $\mu = (\mu(x))_{x \in \mathbb{E}}$ est la loi initiale de la chaîne de Markov. On dit que $\{X_n, n \geq 0\}$ est une chaîne de Markov de matrice de transition P et de loi initiale.

Proposition 2.3

Soit $\{X_n, n \geq 0\}$ une suite de variables aléatoires. $\{X_n, n \geq 0\}$ est une chaîne de Markov de matrice de transition P et de loi initiale μ si et seulement si pour tout $n \in \mathbb{N}$ la loi du vecteur aléatoire $(X_0; \dots; X_n)$ est donnée par :

$$\mathbb{P}(X_0 = x_0; \dots; X_n = x_n) = \mu(x_0) P(x_0; x_1) \dots \mathbb{P}(x_{n-1}; x_n)$$

Démonstration :

$\Rightarrow$ Supposons que $\{X_n, n \geq 0\}$ soit une chaîne de Markov de matrice de transition P et de loi initiale μ . Soit $n \in \mathbb{N}$. On a :

$$\mathbb{P}(X_{0:n} = x_{0:n}) = \mathbb{P}(X_n = x_n \mid X_{0:n-1} = x_{0:n-1}) \mathbb{P}(X_{0:n-1} = x_{0:n-1}) .$$

Puis par définition d'une chaîne de Markov,

$$\mathbb{P}(X_{0:n} = x_{0:n}) = \mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}) \mathbb{P}(X_{0:n-1} = x_{0:n-1}) .$$

Comme P est la matrice de transition de $\{X_n, n \geq 0\}$

$$\mathbb{P}(X_{0:n} = x_{0:n}) = \mathbb{P}(x_{n-1}; x_n) \mathbb{P}(X_{0:n-1} = x_{0:n-1}) .$$

On obtient par itérations :

$$\mathbb{P}(X_{0:n} = x_{0:n}) = \mathbb{P}(x_{n-1}; x_n) \dots \mathbb{P}(x_0; x_1) \mathbb{P}(X_0 = x_0) .$$

Enfin, on conclut en utilisant le fait que μ est la loi initiale de $\{X_n, n \geq 0\}$: $\mathbb{P}(X_{0:n} = x_{0:n}) = \mathbb{P}(x_{n-1}; x_n) \dots \mathbb{P}(x_0; x_1) \mu(x_0)$

$\Leftarrow$) Supposons maintenant que pour tout $n \in \mathbb{N}$ la loi de $(X_0; \dots; X_n)$ est donnée par : $\mathbb{P}(X_0 = x_0; \dots; X_n = x_n) = \mu(x_0) \mathbb{P}(x_0; x_1) \dots \mathbb{P}(x_{n-1}; x_n)$.

Soient $k \in \mathbb{N}$ et $(x_0; \dots; x_{k-1}; x; y) \in \mathbb{E}^{k+2}$

$$\begin{aligned} \mathbb{P}(X_{k+1} = y \mid X_{0:k} = (x_0; \dots; x_k)) &= \frac{\mathbb{P}(X_{0:k+1} = (x_0; \dots; x_k; y))}{\mathbb{P}(X_{0:k} = (x_0; \dots; x_k))} \\ &= \frac{\mu(x_0) \mathbb{P}(x_0; x_1) \dots \mathbb{P}(x_{k-1}; x_k) P(x_k; y)}{\mu(x_0) P(x_0; x_1) \dots P(x_{k-1}; x_k)} \\ &= P(x_k; y) \end{aligned}$$

Donc $\{X_n, n \geq 0\}$ est une chaîne de Markov homogène de matrice de transition P . L'hypothèse nous donne aussi avec $n = 0$ et $x \in \mathbb{E}$, $\mathbb{P}(X_0 = x) = \mu(x)$. Donc $\{X_n, n \geq 0\}$ est une chaîne de Markov de matrice de transition P et de loi initiale μ [3]. ■

2.4 Loi d'une chaîne de Markov

Proposition 2.4

la loi de chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ est entièrement déterminée par la donnée sa matrice de transition p et de la loi de X_0 appelée loi initiale et notée μ_0 pour tout $i \in \mathbb{E}$, $\mu_0(i) := \mathbb{P}(X_0 = i)$. Plus précisément, pour tout entier n et toute suite d'états $i_0, i_1, \dots, i_{n-1}, i_n$ de $\mathbb{E}$.

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \mu_0(i_0) p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}$$

La formule permet d'écrire la probabilité d'une intersection, i.e.

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0)$$

comme un produit de probabilités conditionnelles, les coefficients $p_{i,j}$. En divisant par $\mu_0(i_0)$ dans il s'ensuit que :

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1 \mid X_0 = i_0) = p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}.$$

Démonstration : Le résultat repose sur la formule :

$$\mathbb{P}\left(\bigcap_{k=0}^n A_k\right) = \mathbb{P}(A_0) \mathbb{P}(A_1 \mid A_0) \mathbb{P}(A_2 \mid A_1 \cap A_0) \cdots \mathbb{P}(A_n \mid A_{n-1} \cap \cdots \cap A_0)$$

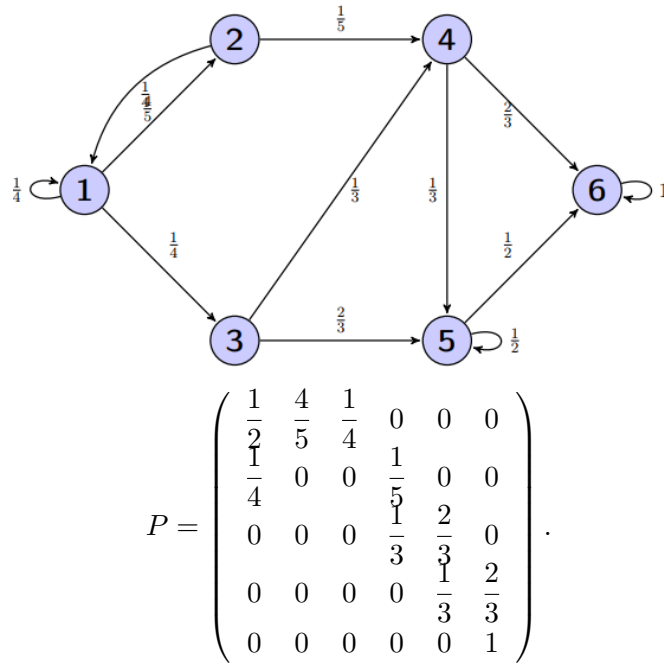
a démontrer par récurrence sur $n \in \mathbb{N}$ et appliquée aux événements $A_k = X_k = i_k$.

Il ne reste plus qu'à identifier les termes : $\mathbb{P}(A_0) = \mathbb{P}(X_0 = i_0)$ vaut $\mu(i_0)$,

$\mathbb{P}(A_1 \mid A_0) = \mathbb{P}(X_1 = i_1 \mid X_0 = i_0)$ vaut par définition p_{i_0, i_1} et

$$\begin{aligned} \mathbb{P}(A_2 \mid A_1 \cap A_0) &= \mathbb{P}(X_2 = i_2 \mid X_1 = i_1, X_0 = i_0) \\ &= \mathbb{P}(X_2 = i_2 \mid X_1 = i_1) \\ &= p_{i_1, i_2} \end{aligned}$$

par la propriété de Markov. C'est pareil pour les autres termes.



Partant de $X_0 = 1$, quel est le chemin le plus probable permettant $X_3 = 6$? On dénombre 3 chemins permettant de rejoindre l'état 6 depuis 1 en 3 coups :

$1 \rightarrow 2 \rightarrow 4 \rightarrow 6$, $1 \rightarrow 3 \rightarrow 4 \rightarrow 6$ et $1 \rightarrow 3 \rightarrow 5 \rightarrow 6$. Leurs probabilités respectives sont :

$$\begin{aligned} \mathbb{P}(X_3 = 6, X_2 = 4, X_1 = 2 \mid X_0 = 1) &= p_{1,2}p_{2,4}p_{4,6} = \frac{1}{2} \times \frac{1}{5} \times \frac{2}{3} = \frac{1}{15} \\ \mathbb{P}(X_3 = 6, X_2 = 4, X_1 = 3 \mid X_0 = 1) &= p_{1,3}p_{3,4}p_{4,6} = \frac{1}{4} \times \frac{1}{3} \times \frac{2}{3} = \frac{1}{18} \\ \mathbb{P}(X_3 = 6, X_2 = 5, X_1 = 3 \mid X_0 = 1) &= p_{1,3}p_{3,5}p_{5,6} = \frac{1}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{12}. \end{aligned}$$

Le plus probable est donc $1 \rightarrow 3 \rightarrow 5 \rightarrow 6$. Dans toute la suite, considérons une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à valeurs dans E, de matrice de transition P et de loi initiale μ_0 . ■

Proposition 2.5

Soient n un entier et i, j des états.

1. $\mathbb{P}(X_n = j \mid X_0 = i) = p_{i,j}^{(n)}$
2. $\mu_{n+1} = \mu_n \mathbb{P}$
3. $\mu_n = \mu_0 \mathbb{P}^n$,

ce qui signifie que $\mathbb{P}(X_n = j)$ est le j élément du vecteur ligne $\mu_0 \mathbb{P}^n$.

En utilisant des vecteurs colonnes pour décrire les lois de X_n et X_{n+1} , le point (2) devient

$${}^t\mu_{n+1} = {}^tP {}^t\mu_n.$$

On peut donc voir l'évolution en loi de la suite $(X_n)_n \in \mathbb{N}$ comme un système itératif

linéaire dont $\mathbb{P}$ est la matrice d'évolution : on obtient la loi de X_{n+1} en multipliant (matriciellement et par la gauche) la loi de X_n par $\mathbb{P}$.

Reprenons l'exemple donné dans la probabilité de rejoindre l'état 6 depuis l'état 1, en 3 coups et quel que soit le chemin emprunté, se calcule :

$$\begin{aligned} p_{1,6}^{(3)} &= \sum_{i,j \in \mathbb{E}} p_{1,i} p_{i,j} p_{j,6} \\ &= p_{1,2} p_{2,4} p_{4,6} + p_{1,3} p_{3,4} p_{4,6} + p_{1,3} p_{3,5} p_{5,6} \\ &= \frac{1}{15} + \frac{1}{18} + \frac{1}{12} \simeq 0.20. \end{aligned}$$

En partant de l'état 1 avec probabilité 1, les lois de X_1 et X_2 se calculent de la manière suivante.

De même, $\mu_5 \simeq (0.15, 0.12, 0.06, 0.04, 0.16, 0.48)$.

Démonstration : Seul le cas $n = 2$ sera traité pour donner l'intuition concernant le point (1). Le coefficient $p_{i,j}^{(2)}$ de la i ligne et de la j colonne de la matrice P^2 vaut

$$p_{i,j}^2 = \sum_{k \in E} p_{i,k} p_{k,j}.$$

La formule des probabilités totales et la proposition permettent d'écrire les égalités :

$$\begin{aligned} \mathbb{P}(X_2 = j, X_0 = i) &= \sum_{k \in E} \mathbb{P}(X_2 = j, X_1 = k, X_0 = i) \\ &= \mu_0(i) \sum_{k \in E} p_{i,k} p_{k,j}, \end{aligned}$$

desquelles on en déduit le résultat ; $\mathbb{P}(X_2 = j \mid X_0 = i) = p_{i,j}^{(2)}$. Par ailleurs

$$\begin{aligned} \mu_{n+1}(j) &= \mathbb{P}(X_{n+1} = j) = \sum_{i \in E} \mathbb{P}(X_{n+1} = j, X_n = i) \\ &= \sum_{i \in E} \mathbb{P}(X_{n+1} = j, X_n = i) \mathbb{P}(X_n = i) \\ &= \sum_{i \in E} p_{i,j} \mathbb{P}(X_n = i). \end{aligned}$$

Cette égalité implique la relation matricielle $\mu_{n+1} = \mu_n \mathbb{P}$ (i.e. le point (2)). Enfin, le point (3) s'obtient par une récurrence immédiate [1]. ■

Proposition 2.6

La loi de la chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ est invariante par translation dans le temps. Autrement dit, pour tous entiers n, m , toute suite d'états $i_0, i_1, \dots, i_n, i_{n+1}, \dots, i_{n+m}$ pour lesquels

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) > 0$$

il vient :

$$\mathbb{P}(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) = \mathbb{P}(X_m = i_{n+m}, \dots, X_1 = i_{n+1} \mid X_0 = i_n).$$

La probabilité conditionnelle d'une trajectoire donnée (i.e. les états $i_{n+1}, \dots, i_{n+m}$) reste la même au cours du temps, qu'elle ait lieu entre les instants $n+1$ et $n+m$ ou entre les instants 1 et m . Seul compte le dernier état visité, en l'occurrence i_n . Dans l'exemple donné dans la gure 1.2, $\text{Fi1} \rightarrow 3 \rightarrow 5 \rightarrow 6$ reste le chemin le plus probable pour rejoindre l'état 6 depuis l'état 1. en 3 coups, que ce soit entre les instants 0 et 3, 47 et 50, ou encore 1234 et 1237 ! Une généralisation de ce résultat est connue sous le nom de relation de Chapman-Kolmogorov :

$$\mathbb{P}(X_{n+m} = j \mid X_0 = i) = \sum_{k \in E} \mathbb{P}(X_m = j \mid X_0 = k) \mathbb{P}(X_n = k \mid X_0 = i).$$

Il faut la lire comme suit : aller de i à j en $n+m$ pas, c'est aller de i à un certain état k en n pas puis de k à j en m pas. Notons par μ_n la loi de X_n . C'est une mesure de probabilité sur E que l'on peut écrire sous la forme d'un vecteur ligne $(\mu_n(j))_{j \in E}$ (i.e. un élément de $R^{\text{Card}(E)}$). L'objectif de la fin de cette section consiste à établir une relation matricielle liant μ_n à la loi initiale μ_0 et à la matrice P . Pour ce faire, notons par $(p(n)_{i,j})_{i,j \in E}$ les coefficients de la matrice P^n , puissance n de P . L'expression brute de $p_{i,j}^{(n)}$ est

$$p_{i,j}^{(n)} = \sum_{i_1, \dots, i_{n-1} \in E} p_{i,i_1} p_{i_1,i_2} \dots p_{i_{n-1},j}$$

Cette formule étant purement algébrique, voici une nouvelle expression du coefficient $p_{i,j}^{(n)}$ lui donnant davantage de sens.

2.5 Chaînes de Markov sur un ensemble fini

Chaîne de markov irréductibles

Définition 2.8

Soient i et j deux états de E . On dit que l'état j est accessible depuis l'état i si $\exists n \in \mathbb{N}$ $p_{i,j}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i) > 0$.

On dit que les états i et j communiquent si chacun est accessible depuis l'autre.

On note alors $i \longleftrightarrow j$.

Proposition 2.7

La relation $\longleftrightarrow$ est une relation d'équivalence sur E .

Démonstration : La réflexivité (i.e. $i \longleftrightarrow i$) est immédiate :

pour $n = 0$, $\mathbb{P}(X_0 = i \mid X_0 = i) = 1$. Il en va de même pour la symétrie ($i \longleftrightarrow j$ implique $j \longleftrightarrow i$). Enfin, la transitivité repose sur la relation de Chapman- Kolmogorov. Supposons que $i \longleftrightarrow j$ et $j \longleftrightarrow k$. En particulier, les états i et j sont respectivement accessibles depuis i et j . Il existe donc des entiers m et n tels que $\mathbb{P}(X_m = k \mid X_0 = j) > 0$ et $\mathbb{P}(X_n = j \mid X_0 = i) > 0$. Par conséquent,

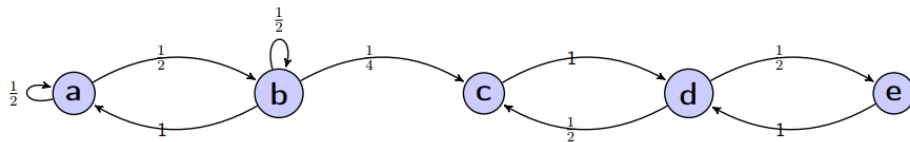
$$\begin{aligned} \mathbb{P}(X_{n+m} = k \mid X_0 = i) &= \sum_{l \in E} \mathbb{P}(X_m = k \mid X_0 = l) \mathbb{P}(X_n = l \mid X_0 = i) \\ &\geq \mathbb{P}(X_m = k \mid X_0 = j) \mathbb{P}(X_n = j \mid X_0 = i) > 0 \end{aligned}$$

d'où k est accessible depuis l'état i . C'est la même chose dans l'autre sens. L'espace E peut donc être partitionné en classes d'équivalence pour la relation $\longleftrightarrow$, appelées classes irréductibles. Nous insisterons dans les paragraphes suivants sur le fait que les états d'une même classe irréductible ont des propriétés équivalentes vis à vis de la chaîne (récence, transience et périodicité). Lorsque l'espace E est réduit à une seule classe (i.e. tous les états communiquent), on dit que la chaîne est irréductible. En général, E se partitionne en états isolés dans lesquels on ne revient jamais une fois qu'on les a quittés, et en classes irréductibles disjointes. Pour déterminer les classes irréductibles d'une chaîne de Markov, il est commode de travailler sur le graphe de transition plutôt que sur la matrice de transition P . ■

Exemple 2.2

Considérons une chaîne de Markov à valeurs dans $E = (a, b, c, d, e)$ et dont la matrice et le graphe de transition sont données par :

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$



La chaîne comporte deux classes irréductibles : (a, b) et (c, d, e) [1].

Chaînes de Markov régulières

Définition 2.9

soit la matrice de transition

$$P = \begin{pmatrix} w_1 & w_2 & w_3 & \dots & \dots & w_m \\ w_1 & w_2 & w_3 & \dots & \dots & w_m \\ w_1 & w_2 & w_3 & \dots & \dots & w_m \\ w_1 & w_2 & w_3 & \dots & \dots & w_m \end{pmatrix}$$

est dit régulière si la matrice w^2 est positif.

Exemple 2.3

Considérons la matrice de transition

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

La chaîne est-elle régulière ? la matrice

$$P^2 = P \cdot P = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

les états de P^2 sont positifs, donc la chaîne est régulière [2].

Chaînes de Markov réversibles

Dans cette section, l'ensemble X peut être infini (mais doit être dénombrable). Rappelons que E^X désigne l'ensemble des applications $f : X \rightarrow E$.

Définition 2.10

Soit P une matrice stochastique. Un vecteur $a = a_{i \in X} \in [0, \infty)^X, a \neq 0$, est dit réversible par rapport à P si $a_i p_{ij} = a_j p_{ji} \forall i, j \in X$. Une chaîne de Markov est dite réversible si une matrice admet un vecteur réversible.

La condition est appelée condition d'équilibre détaillé en physique. Elle signifie que si

les états i et j sont occupés avec probabilités proportionnelles à a_i et a_j respectivement, alors les taux de transition de i à j et de j à i sont égaux.

Théorème 2.1

Soit P une matrice stochastique et $a \in [0, \infty)^X$ un vecteur non nul.

1. Si a est réversible par rapport à P , alors a est une mesure invariante.
2. Si a est réversible par rapport à P , et $\sum_{j \in X} a_j < \infty$, alors la mesure π définie par

$$\pi_i = a_i / \sum_{j \in X} a_j \text{ est une distribution stationnaire.}$$

3. Si π est une distribution stationnaire, alors

$$P_\pi X_0 = i_0, X_1 = i_1, \dots, X_n = i_n = P_\pi X_0 = i_n, X_1 = i_{n-1}, \dots, X_n = i_0$$

pour tout $n \in \mathbb{N}$ et tout choix de $i_0, \dots, i_n \in X$.

Démonstration :

1. On a X

$$\sum_{i \in X} a_i p_{ij} = a_j \sum_{i \in X} p_{ji} = a_j$$

2. Suit immédiatement de 1.
3. Par le Théorème

$$\begin{aligned} P_\pi X_0 = i_0, X_1 = i_1, \dots, X_n = i_n &= \pi_{i_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n} \\ &= p_{i_1 i_0} \pi_{i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n} \end{aligned}$$

qui est égal à $P_\pi X_0 = i_n, X_1 = i_{n-1}, \dots, X_n = i_0$.

La relation signifie qu'une trajectoire a la même probabilité que la trajectoire renversée dans le temps. C'est ce qui justifie le terme de réversibilité. ■

Exemple 2.4

On considère une chaîne de Markov d'espace d'états $E := \{1; 2; 3\}$ et de matrice de transition

$$p = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}$$

la chaîne est-elle réversible ?

la chaîne n'est pas réversible car : on donne

$$\pi = \frac{1}{9}(4; 2; 3)$$

donc

$$\pi_1 p_{12} = \frac{2}{9} \neq \pi_2 p_{21} = \frac{1}{9}$$

donc n'est pas réversible [4].

Chaînes de Markov absorbantes

Définition 2.11

On dit qu'un état $j \in E$ est accessible depuis un autre état $i \in E$, et on note $i \rightsquigarrow j$, s'il existe un temps $n \in \mathbb{N}$ tel que $p_{ij}^{(n)} > 0$, c'est-à-dire que partant de i , on atteint j avec probabilité positive en un nombre fini de pas. On notera $i \sim j$ si on a à la fois $i \rightsquigarrow j$ et $j \rightsquigarrow i$. On vérifie facilement que la relation est réflexive et transitive, et que $\sim$ est une relation d'équivalence [?].

Définition 2.12

Un état $i \in X$ est dit absorbant si $p_{ii} = 1$ (et donc nécessairement $p_{ij} = 0$ pour tout $j \neq i$). Une chaîne de Markov est dite absorbante s'il existe, pour tout état de X , un état absorbant accessible depuis cet état. Dans le reste de cette section, nous allons considérer des chaînes absorbantes avec $r \geq 1$ états absorbants. Nous conviendrons de numérotter les états de manière à placer d'abord les $q = N - r$ états non absorbants, et ensuite les r états absorbants. La matrice de transition prend alors la forme canonique

$$p = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix}$$

où Q est une matrice de taille $q \times q$, R est une matrice de taille $q \times r$, 0 désigne la matrice nulle de taille $r \times q$, et I la matrice identité de taille r . Il est facile de montrer par récurrence que

$$p^n = \begin{pmatrix} Q^n & [I + Q + \dots + Q^{n-1}] R \\ 0 & I \end{pmatrix}$$

Proposition 2.8

Soit P la matrice de transition d'une chaîne de Markov absorbante, écrite sous forme canonique. Alors

1. On a

$$\lim_{n \rightarrow +\infty} Q^n = 0.$$

2. La matrice $I - Q$ est inversible, et son inverse vaut

$$[I - Q]^{-1} = \lim_{n \rightarrow \infty} Q^n.$$

Démonstration :

1. Soit $i \neq q$ un état non absorbant. L'élément de matrice $(Q^n)_{ij}$ de Q^n est la probabilité de se trouver dans l'état non absorbant j , après n pas, partant de i . Par conséquent, $(Q^n)_{ij}$ est inférieur ou égal à la probabilité de ne pas avoir atteint d'état absorbant en n pas. Soit

$$m_i = \min \{n \geq 1 : \exists k > q, (P^n)_{ik} > 0\}$$

le nombre minimal de pas nécessaire à atteindre un état absorbant k depuis i . Soit

$$p_i = P_i \{X_{m_i} \leq q\} < 1$$

la probabilité de ne pas atteindre d'état absorbant en m_i pas, partant de i . Soit enfin

$$M = \max_{i=1, \dots, q} m_i$$

$p = \max_{i=1} p_i$. Alors la probabilité de ne pas atteindre d'état absorbant en M pas, par tant de n'importe quel état non absorbant i , est bornée par p^M . Cette probabilité tend vers 0 lorsque M tend vers l'infini. La probabilité de ne pas être absorbé après un nombre arbitraire m de pas étant une fonction décroissante de m , elle tend nécessairement vers 0. Par conséquent, $(Q^n)_{ij}$ tend vers zéro lorsque n tend vers l'infini, pour tout $j \in \{1, \dots, q\}$.

2. Supposons qu'il existe un vecteur x tel que $Qx = x$. Dans ce cas on a

$$x = Qx = Q^2x = \dots = Q^nx = \dots = \lim_{n \rightarrow \infty} Q^nx = 0$$

ce qui montre que Q n'admet pas la valeur propre 1. Par conséquent, $I - Q$ est inversible. Enfin, comme

$$[I - Q] \sum_{k=0}^n Q^k = I - Q^{n+1} \rightarrow I$$

lorsque $n \rightarrow \infty$ on obtient multipliant à gauche par $[I - Q]^{-1}$. Nous noterons F la matrice $[I - Q]^{-1}$, et nous l'appellerons la matrice fondamentale de la chaîne. montre que

$$\lim_{n \rightarrow \infty} Q^n = \begin{pmatrix} 0 & FR \\ 0 & I \end{pmatrix}$$

Le fait que Q^n tend vers zéro traduit donc le fait que la probabilité d'absorption tend vers 1 lorsque le temps tend vers l'infini. La matrice $B = FR$ devrait représenter les probabilités de transition, dans la limite des temps infinis, entre états non absorbants et absorbants. Ceci est confirmé par le résultat suivant ■

Exemple 2.5 (*Jeu de Pile ou Face*)

Anatole et Barnabé jouent 'a la variante suivante de Pile ou Face. Ils jettent une pièce de monnaie (parfaitement équilibrée) de manière répétée.

Anatole gagne dès que la pièce tombe trois fois de suite sur Face, alors que Barnabé gagne dès que la suite Pile-Face-Pile apparait.

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

$$R = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \\ 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix}$$

On calcule alors la matrice fondamentale

$$F = [I - Q]^{-1} = \begin{pmatrix} \frac{7}{3} & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{3}{4} & \frac{1}{3} & \frac{3}{2} \\ \frac{3}{4} & \frac{3}{4} & \frac{3}{4} & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} & \frac{3}{2} & \frac{3}{4} \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$$B = FR = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{3}{2} \\ \frac{3}{4} & \frac{3}{2} \\ \frac{1}{3} & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Ainsi, partant de l'un des états PP, PF ou FP , Anatole gagne avec probabilité $1/3$, et Barnabé gagne avec probabilité $2/3$. Partant de l'état FF , c'est Barnabé qui gagne avec probabilité $1/3$, et Anatole qui gagne avec probabilité $2/3$. Comme personne ne gagne lors des deux premiers jets, et que les quatre états PP, PF, FP et FF sont atteints avec la même probabilité, il faut choisir la distribution initiale $v = \left(\frac{1}{4}; \frac{1}{4}; \frac{1}{4}; \frac{1}{4}; 0, 0\right)$. Par conséquent, Anatole gagne le jeu avec probabilité

$$P_v \{X_\tau = \text{"Agagne"}\} = \sum_{i=1}^4 v_i b_{i1} = \frac{5}{12}.$$

[4].

Théorème 2.2

Soit F la matrice fondamentale d'une chaîne de Markov absorbante.

1. L'élément de matrice f_{ij} de F est l'espérance du nombre de passages en j partant de i :

$$f_{ij} = E_i \sum_{n \geq 0} \mathbb{1}_{\{X_n = j\}}.$$

2. Soit $\tau = \inf \{n \geq 1 : X_n > q\}$ la variable aléatoire donnant le temps jusqu'à absorption. Alors

$$E_i(\tau) = \sum_{j=1}^q f_{ij}.$$

3. Les éléments de matrice b_{ik} de $B = FR$ donnent les probabilités d'être absorbés dans les différents états :

$$b_{ik} = P \{X_\tau = k\}.$$

2.6 Chaînes de Markov sur un ensemble dénombrable

Récurrence et transition

Définition 2.13

Soit $i \in E$. La v.a. $\mathbb{T}_i$ définie par

$$\mathbb{T}_i = \inf \{n \geq 1, X_n = i\}$$

est appelée temps d'atteinte de i ou encore temps de retour à i lorsque la chaîne $(X_n)_{n \in \mathbb{N}}$ part de i . Par convention, lorsque pour tout $n \geq 1$, $X_n \neq i$, on pose $\mathbb{T}_i = +\infty$.

Il faut bien comprendre que l'événement

$$\{\mathbb{T}_{i < +\infty}\} = \{\exists n \geq 1, X_n = i\} = \cup_{n \geq 1} \{X_n = i\}$$

signifie que la chaîne $(X_n)_{n \in \mathbb{N}}$ repassera par l'état i (au moins une fois) alors que

$$\{\mathbb{T}_i = +\infty\} = \{\forall n \geq 1, X_n \neq i\} = \cap_{n \geq 1} \{X_n \neq i\}$$

signifie que la chaîne $(X_n)_{n \in \mathbb{N}}$ ne repassera jamais par i .

Définition 2.14

Un état $i \in E$ est dit récurrent si, partant de i , on y revient presque sûrement en temps fini.

$$\mathbb{P}(\mathbb{T}_{i<+\infty} \mid X_0 = i) = 1$$

L'état i est dit transient dans le cas contraire, i.e. lorsque $\mathbb{P}(\mathbb{T}_{i<+\infty} \mid X_0 = i) > 0$. Autrement dit, un état est transient si avec probabilité strictement positive, on peut le quitter pour ne jamais y revenir. Comme cas particulier d'état transient, on retrouve les états pour lesquels $p_{i,i}^{(n)} = 0$ pour tout $n \geq 1$. Ce sont ceux que l'on quitte au premier pas pour ne jamais y revenir :

$$\begin{aligned} \mathbb{P}(\mathbb{T}_{i<+\infty} \mid X_0 = i) &= \mathbb{P}(\exists \geq 1, X_n = i \mid X_0 = i) \\ &= \mathbb{P}(\cup_{n \geq 1} \{X_n = i\} \mid X_0 = i) \\ &\leq \sum_{n \geq 1} \mathbb{P}(X_n = i \mid X_0 = i) \\ &\leq \sum_{n \geq 1} p_{i,i}^{(n)}. \end{aligned}$$

Lemme 2.1

Pour tout entier n et tous états i et j

$$\mathbb{P}(N_i \geq n + 1 \mid X_0 = j) = \mathbb{P}(\mathbb{T}_i < +\infty \mid X_0 = j) \mathbb{P}(N_i \geq n \mid X_0 = i).$$

Les états transients sont ceux dans lesquels on ne passe qu'un nombre fini de fois. Par opposition, on revient une infinité de fois dans un état récurrent.

Proposition 2.9

Les conditions suivantes sont équivalentes :

1. l'état i est récurrent : $\mathbb{P}(\mathbb{T}_i < +\infty \mid X_0 = i) = 1$;
2. conditionnellement à $X_0 = i$, la chaîne de markove $(X_n)_{n \in \mathbb{N}}$ revient presque sûrement une infinité de fois en i :

$$\mathbb{P}(N_i = +\infty \mid X_0 = i) = 1.$$

3. la série $\sum p_{i,i}^{(n)}$ diverge. Les conditions suivantes sont équivalentes :
4. l'état i est transient : $\mathbb{P}(\mathbb{T}_i < +\infty \mid X_0 = i) < 1$.
5. conditionnellement à $X_0 = i$, la v.a. $N_i + 1$ suit la loi géométrique de paramètre $1 - a_i$, avec $a_i = \mathbb{P}(\mathbb{T}_i < +\infty \mid X_0 = i)$:

$$\forall n \in \mathbb{N}, \mathbb{P}(N_i = n \mid X_0 = i) = a_i^n (1 - a_i).$$

(en particulier N_i est presque sûrement finie).

6. conditionnellement à $X_0 = i$, la variable aléatoire N_i est intégrable : $E[N_i | X_0 = i] = \sum p_{i,i}^{(n)} < +\infty$ Remarquons que l'identité entre l'espérance conditionnelle $E[N_i | X_0 = i]$ et la série $p_{i,i}^{(n)}$ est valable quelle que soit la nature de cette dernière, en la considérant comme un élément de $[0, +\infty]$. Elle repose sur le théorème de Fubini.

$$\begin{aligned} \mathbb{E}[N_i | X_0 = i] &= \mathbb{E}\left[\sum_{n \geq 1} \mathbb{1}_{X_n = i} | X_0 = i\right] \\ &= \sum_{n \geq 1} \mathbb{E}[\mathbb{1}_{X_n = i} | X_0 = i] \\ &= \sum_{n \geq 1} \mathbb{P}[\mathbb{1}_{X_n = i} | X_0 = i] \\ &= \sum_{n \geq 1} p_{i,i}^{(n)}. \end{aligned}$$

Reprenons l'exemple On a déjà établi que

$$\mathbb{P}(T_b = +\infty | X_0 = b) \geq \frac{1}{4}.$$

On peut montrer qu'il y a en fait égalité (en effet, partir de b pour aller en a et ne jamais en revenir est un événement non vide, certes, mais de probabilité nulle). Ainsi la Proposition affirme que le nombre de passages en l'état b , partant de b , suit la loi géométrique de paramètre $\frac{3}{4}$. On en déduit son espérance :

$$\mathbb{E}[N_b | X_0 = b] = \mathbb{E}[N_b + 1 | X_0 = b] - 1 = \frac{1}{1 - \frac{3}{4}} - 1 = 3.$$

Les deux propositions ci-dessus se démontrent simultanément.

Démonstration :

Prouvons (1) $\Rightarrow$ (2). En combinant, $\mathbb{P}(T_i < +\infty | X_0 = i) = 1$ et le Lemme, il vient

$$\mathbb{P}(N_i \geq n + 1 | X_0 = i) = \mathbb{P}(N_i \geq n | X_0 = i) = \dots = \mathbb{P}(N_i \geq 1 | X_0 = i),$$

en itérant. Par ailleurs, $N_i \geq 1 = \{t_i < +\infty\}$. On en déduit que pour tout n , $\mathbb{P}(N_i \geq n | X_0 = i) = 1$ Il ne reste plus qu'à faire tendre n vers l'infini pour obtenir (2) :

$$\begin{aligned} \mathbb{P}(N_i = +\infty | X_0 = i) &= \mathbb{P}(\cap_{n \geq 1} \{N_i \geq n\} | X_0 = i) \\ &= \lim_{n \rightarrow +\infty} \mathbb{P}(N_i \geq n | X_0 = i) \\ &= 1. \end{aligned}$$

L'implication (2) $\Rightarrow$ (3) repose sur :

$$\sum_{n \geq 1} p_{i,i}^{(n)} = \mathbb{E}[N_i | X_0 = i] \geq (+\infty) \mathbb{P}(N_i = +\infty | X_0 = i) = +\infty.$$

Supposons (4) : $a_i = \mathbb{P}(T_i < +\infty \mid X_0 = i) < 1$. Le Lemme donne :

$$\mathbb{P}(N_i \geq n+1 \mid X_0 = i) = a_i \mathbb{P}(N_i \geq n \mid X_0 = i) = \dots = a_i^n \mathbb{P}(N_i \geq 1 \mid X_0 = i).$$

Ainsi,

$$P(N_i \geq n+1 \mid X_0 = i) = a_i^n P(T_i < +\infty \mid X_0 = i) = a_i^{n+1}.$$

On en déduit d'une part que N_i est presque sûrement finie

$$P(N_i = +\infty \mid X_0 = i) = \lim_{n \rightarrow +\infty} P(N_i \geq n \mid X_0 = i) = 0$$

(car $a_i < 1$) et d'autre part que sa loi est géométrique de paramètre a_i . Ce qui prouve (5).

$$P(N_i = n \mid X_0 = i) = P(N_i \geq n \mid X_0 = i) - P(N_i \geq n+1 \mid X_0 = i) = a_i^n (1 - a_i).$$

L'implication (5) $\Rightarrow$ (6) est immédiate :

$$E[N_i \mid X_0 = i] = \sum_{n \geq 1} P(N_i \geq n \mid X_0 = i) = \frac{a_i}{1 - a_i} < +\infty$$

Enfin, (3) $\Rightarrow$ (1) et (6) $\Rightarrow$ (4) sont respectivement les contraposées de (4) $\Rightarrow$ (6) et (1) $\Rightarrow$ (3).

Les boucles sont bouclées ! [1] ■

Exemple 2.6

On considère une chaîne de Markov d'espace d'états $E := \{1; 2; 3; 4\}$ et de matrice de transition

$$p = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1. déterminer les états transitoires et récurrents de la chaîne ; Les classes de la chaîne sont les suivantes : $\{1\}$, $\{2; 3\}$ et $\{4\}$. Les classes $\{1\}$ et $\{4\}$ sont clairement récurrentes puisque pour $x \in \{1; 4\}$ on a $P_x(T_x < \infty) = 1$. La classe $\{2; 3\}$ est quant à elle transiente. En effet, si elle était récurrente,
2. serait récurrent et on aurait $N_2 = +\infty P_2 - p$. où $N_2 = \sum_{n \geq 0} 1_{X_n} = 2$. Mais $N_2 = 1$ sur l'évènement $\{X_1 = 1\}$ qui est de mesure > 0 puisque $P_2(X_1 = 1) = p(2; 1) = \frac{1}{2} > 0$ [4].

Les chaînes dans tous leurs états

Soit (X_n) une $CM(E, \nu, P)$. Définissons la quantité $\rho(x, y) := P^x(T_y < 1)$. Un état x est qualifié de récurrent lorsque les propriétés équivalentes suivantes sont réalisées : $\rho(x, x) = 1, G(x, x) = \infty, H(x, x) = \infty$. Il est qualifié de transitoire sinon. Le tableau 1 regroupe des propriétés utiles. Pour un état récurrent x , la propriété de Markov forte appliquée récursivement permet de définir une infinité de temps de retours successifs en x , qui délimitent des excursions de même loi enracinées en x (cela correspond à un système de renouvellement sur N^*). On note E_R l'ensemble des états récurrents et $E_T = E \setminus E_R$ celui des états transitoires. On dit que $F \subset E$ est clos lorsque $P^x(T_{F^c} < \infty) = 0$ pour tout $x \in F$. La chaîne ne s'échappe jamais d'un ensemble clos. Un état x est dit absorbant (ou point cimetière) lorsque les conditions équivalentes suivantes sont réalisées : x est clos, $P(x, x) = 1$.

État x récurrent	Etat x transitoire
$\rho(x, x) = 1, H(x, x) = \infty$	$\rho(x, x) < 1, H(x, x) < \infty$
$\rho(x, y) > 0 \Rightarrow P^x(N_y = \infty) = 1$	$P^x(N_x = \infty) = 0$
$\rho(y, x) = P^y(N_x = \infty)$	$L(N_x X_0 = x) = G(\rho(x, x))$
$\rho(y, x) > 0 \Rightarrow H(y, x) = \infty$	$H(y, x) = \rho(y, x)/(1 - \rho(x, x)) < \infty$

Exemple 2.7

(Processus de Bernoulli).

Considérons le processus de Bernoulli

$$p(x, y) = \begin{cases} p & \text{si } y = x + 1 \\ 1 - p & \text{si } y = x \\ 0 & \text{sinon} \end{cases}$$

Si $p > 0$, alors tous les états sont transitoires car $P^x(T_x = \infty) = 1$. Si $p = 0$, tous les états sont absorbants (la chaîne est constante) [9].

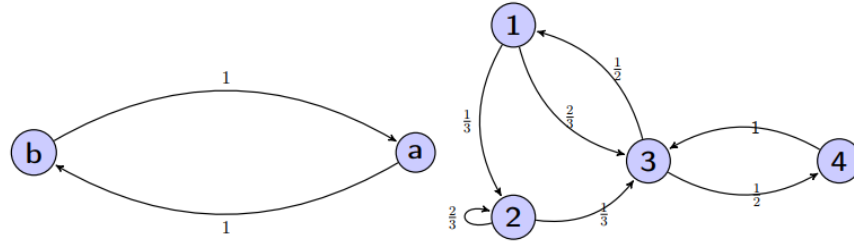
Périodicité

Définition 2.15

La période d'un état i est l'entier $d(i)$ défini par $d(i) = \text{PGCD } n \geq 1, p_{i,i}^{(n)} > 0$. Lorsque $d(i) = 1$, l'état i est qualifié de apériodique. Un état en lequel on peut rester avec probabilité non nulle, i.e. $p_{i,i} > 0$, est automatiquement apériodique. Voici deux exemples. à gauche, l'état a est de période 2 (idem pour b). En effet, chaque chemin de probabilité non nulle partant de a et y revenant comprend un nombre pair de pas. à droite, tous les états sont apériodiques. étudions en particulier le cas de l'état 1. Depuis cet état, il est possible d'y revenir en 3 pas (en faisant $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$) :

$$p_{1,1}^{(3)} = \sum_{i,j \in E} p_{1,i} p_{i,j} p_{j,1} \geq p_{1,2} p_{2,3} p_{3,1} > 0$$

Il est également possible d'y revenir en 5 pas : $p_{1,1}^{(5)} > 0$ (en faisant $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 3 \rightarrow 1$). Puisque 3 et 5 sont premiers entre eux, l'état 1 est apériodique.



La périodicité est une propriété de classe irréductible.

Proposition 2.10

Si les états i et j communiquent alors ils ont la même période. Tous les états d'une même classe irréductible ont donc la même période. Si celle-ci vaut 1, la classe est alors qualifiée de apériodique.

Démonstration : Soient i et j deux états qui communiquent. Il suffit de montrer que $d(j)$ divise $d(i)$. En effet, par symétrie, on aura également $d(i)$ divise $d(j)$, et donc $d(j) = d(i)$. Comme i et j communiquent, il existe deux entiers 1 et m tels que $p_{i,j}^{(l)} > 0$ et $p_{j,i}^{(m)} > 0$. Considérons maintenant un entier n tel que $p_{i,i}^{(n)} > 0$. Les inégalités

$$p_{j,j}^{(m+n+l)} \geq p_{j,i}^{(m)} p_{i,i}^{(n)} p_{i,j}^{(l)} > 0$$

$$et p_{j,j}^{(m+l)} \geq p_{j,i}^{(m)} p_{i,j}^{(l)} > 0$$

impliquent respectivement que $d(j)$ divise les entiers $m+n+l$ et $m+l$: il divise donc la différence, i.e. n . Autrement dit, $d(j)$ divise tous les entiers n tels que $p_{i,i}^{(n)} > 0$. Il divise donc leur PGCD $d(i)$ [1]. ■

Exemple 2.8

On considère 5 points équirépartis sur un cercle. Un promeneur saute à chaque instant, d'un point à l'un de ses voisins avec la probabilité $1/2$ pour chaque voisin. Déterminer la matrice de transition de la chaîne ainsi obtenue. Quelle est la période de ses états ? L'ensemble des états est $E = \{1, 2, 3, 4, 5\}$ la matrice de transition

$$p = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

La chaîne est irréductible (tous les états communiquent) et finie donc récurrente positive. Comme tous les états sont dans la même classe, leur période est la même. On a

$$d(1) = \text{PGCD} \{n \geq 1; p_{1,1}^{(n)} > 0\}.$$

Or

$$p_{1,1}^{(2)} \geq p_{1,2}p_{2,1} = \frac{1}{4} > 0$$

et $p_{1,1}^{(5)} \geq p_{1,2}p_{2,3}p_{3,4}p_{4,5}p_{5,1} = \frac{1}{2^5} > 0$. Or le seul diviseur commun de 2 et 5 est 1. On a donc $d = d(1) = 1$ cette chaîne est donc aperiodique [4].

Distributions stationnaires

Définition 2.16

Supposons que $\rho < 1$, et soit $\Pi = (\pi_0; \pi_1; \pi_2; \dots)$ la distribution de la chaîne de Markov induite. Il ne sera généralement pas possible de trouver la distribution π elle-même, mais nous pouvons calculer la fonction génératrice correspondante Π

$$\begin{aligned}\Pi &= \Pi P \\ \pi_j &= \sum_{i=0}^{\infty} p_{ij} \pi_i \quad i = 1; 2; \dots \\ \pi_j &= a_j \pi_0 + \sum_{i=1}^{j+1} a_{j-i+1} \pi_i\end{aligned}$$

$$\pi_j = a_j \pi_0 + \sum_{i=0}^{j+1} a_{j-i+1} \pi_i - a_{j+1} \pi_0; \quad (j = 0; 1; 2; \dots)$$

Si l'on multiplie cette équation précédente par z^j et si l'on somme sur j , on a :

$$\begin{aligned}\sum_{j=0}^{\infty} z_j \pi_j &= \pi_0 \sum_{j=0}^{\infty} z_j a_j + \sum_{j=0}^{\infty} z_j \sum_{i=0}^{j+1} a_{j-i+1} \pi_i - \pi_0 \sum_{j=0}^{\infty} z_j a_{j+1} \\ &= \pi_0 \sum_{j=0}^{\infty} z_j a_j + \frac{1}{z} \sum_{j=0}^{\infty} c_j + 1 z^{j+1} - \frac{\pi_0}{z} \sum_{j=0}^{\infty} z^{j+1} a_{j+1}\end{aligned}$$

ou' $c_{j+1} = \sum_{i=0}^{j+1} a_{j-i+1} \pi_i$ est le terme général du produit de convolution des deux distributions $a = (a_0; a_1; a_2; \dots)$ et Π . En introduisant les fonctions génératrices

$$\begin{aligned}\Pi(z) &= \sum_{i=0}^{\infty} \pi_i Z^i \\ a(z) &= \sum_{i=0}^{\infty} a_i Z^i\end{aligned}$$

et

$$cz() = \sum_{i=0}^{\infty} c_i Z^i = \Pi(z) a(z)$$

on obtient

$$\Pi(z) = \pi_0 a_z + \frac{1}{z} [c(z) - c_0] - \frac{\pi_0}{z} [a(z) - a_0]$$

d'ou'

$$\Pi(z) = \frac{\pi_0 a(z) (z - 1)}{z - a(z)}$$

alors

$$E(a_{n+1}) = P(X_n > 0) = 1 - P(X_n = 0)$$

d'ou'

$$\rho = 1 - \pi_0) \Rightarrow \pi_0 = 1 - \rho$$

et la fonction génératrice de la distribution stationnaire s'écrit

$$\Pi(z) = \frac{(1 - \rho)a(z)(z - 1)}{z - a(z)}$$

[5].

Exemple 2.9

On considérons la matrice suivants :

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

La distribution stationnaire :

d'pres la forme générale

$$\pi = \pi.p$$

on applique la formule

$$(\pi_1, \pi_2, \pi_3, \pi_4) \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix} = (\pi_1, \pi_2, \pi_3, \pi_4)$$

$$\begin{cases} \frac{\pi_2}{4} = \pi_1 & (1) \\ \pi_1 + \frac{\pi_2}{2} + \frac{\pi_3}{2} = \pi_2 & (2) \\ \frac{\pi_2}{2} + \pi_4 = \pi_3 & (3) \\ \frac{\pi_3}{2} = \pi_4 & (4) \end{cases}$$

on remplace (1) et (4) dans (2) et (3) on trouve

$$\begin{cases} \pi_1 + 2\pi_1 + \frac{\pi_3}{2} = 4\pi_1 & (5) \\ 2\pi_1 + \frac{\pi_3}{2} = \pi_3 & (6) \end{cases}$$

on trouve

$$\begin{cases} \pi_1 = \pi_4 \\ \pi_2 = 4\pi_1 \\ \pi_3 = 2\pi_1 \end{cases}$$

la distribution stationnaire $\pi = (\pi_1, \pi_2, \pi_3, \pi_4) = \left(\frac{1}{8}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}\right)$ [4].

2.7 Chaînes de Markov à temps discret

Définition 2.17

on appelle chaîne de Markov à temps discret un processus stochastique à espace d'état discret et à temps discret et qui vérifie la propriété d'absence de mémoire c-à-d : Un processus stochastique $\{X_n\}_{n \in \mathbb{N}}$ à valeurs dans l'espace mesurable (E, ϵ) est markovien si et seulement s'il vérifie la propriété de Markov

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

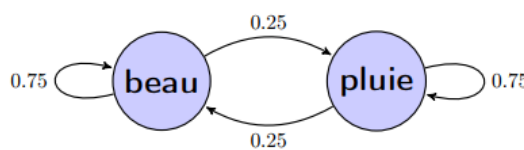
pour tout $n \in \mathbb{N}$, pour tout état j et pour toute suite d'états $i_0, \dots, i_{n-1}$, i pour lesquels la probabilité conditionnelle à un sens. On peut alors définir la probabilité de transition d'un état i vers un état j , par p_{ij}

$$p_{ij} = P(X_n = j \mid X_{n-1} = i), \quad \forall n \in \mathbb{N}$$

La matrice de transition $P = [p_{ij}]_{i,j \in E}$ est une matrice carrée d'ordre fini ou infini [5].

Exemple 2.10

On dispose de statistiques, de tableaux de chances, indiquant qu'il fait beau 50 façon naïve de prédire le temps est la suivante $P(\text{beau}) = P(\text{mauvais}) = 0.5$. La justification est la loi des grands nombres : l'hypothèse de base est que le climat du lendemain est indépendant du climat des autres jours. Malheureusement, on constate en examinant les chiffres qu'il y a 3 fois plus de chances que le climat du lendemain reste le même que celui d'aujourd'hui (plutôt qu'il ne change). On représente ce modèle par le diagramme



ou encore par la matrice

$$P = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$$

2.8 Chaînes de Markov à temps continu

On s'intéresse à des processus $(X_t)_{t \geq 0}$ indexés par le temps continu $t \in R^+$ à valeurs dans un espace fini ou dénombrable I . Une trajectoire du processus est donc une fonction $t \rightarrow X_t(w)$ de R_+ dans I qui dépend d'un aléa w , tout comme une trajectoire d'une chaîne de Markov $(Y_n)_{n \in N}$ est une suite $(Y_n(w))_{n \in N}$ à valeurs dans I qui dépend de l'aléa w . On supposera dans toute la suite que les trajectoires de X sont continues à droite : pour (presque tout) w et tout $t \geq 0$, il existe $\epsilon \geq 0$ tel que

$$\forall s \in [t, t + \epsilon], X_s(w) = X_t(w).$$

Pour une trajectoire donnée, trois comportements différents sont alors possibles :

1. la trajectoire devient constante à partir d'un certain instant,
2. la trajectoire possède de un nombre infini de sauts sur R^+ mais seulement un nombre fini sur tout intervalle compact,
3. la trajectoire possède de un nombre infini de sauts dans un intervalle de temps compact. Il existe alors un temps d'explosion ζ fini. Après ce dernier instant d'explosion, la trajectoire peut repartir, exploser à nouveau ou non... Dans toute la suite, nous considérons les processus jusqu'à leur premier temps d'explosion. On parle alors de processus minimal. On appelle $J_0, J_1, \dots$ les temps de saut de $(X_t)_{t \geq 0}$ et $S_1, S_2, \dots$ les temps d'attente. Ils se déduisent de $(X_t)_{t \geq 0}$ par les relations

$$J_0 = 0, J_{n+1} = \inf \{t \geq J_n; X_t \neq X_{J_n}\}$$

pour $n \in N$ avec la convention $\inf \emptyset = +\infty$ et pour $n \in N^*$.

$$S_n = \begin{cases} J_n - J_{n-1} & \text{si } n \geq 1 \\ \infty & \text{sinon} \end{cases}$$

Remarque 2.3

La propriété de continuité à droite assure que $S_n > 0$ pour tout n . Si $J_{n+1} = \infty$ alors on peut définir $X_1 = X_{J_n}$, sinon, X_1 n'est pas défini. Le (premier) temps d'explosion ζ est défini par

$$\zeta = \sup_n J_n = \sum_{n=1}^{\infty} S_n$$

Le processus à temps discret $(Y_n)_{n \in N}$ défini par $Y_n = X_{J_n}$ pour tout $n \in N$ est appelée chaîne incluse. Il représente la suite des valeurs prises par $(X_t)_{t \geq 0}$.

Définition 2.18

Soit I un espace dénombrable. Un générateur sur I est une matrice $Q = (q_{ij})_{i,j \in I}$ vérifiant

1. pour tout $i \in I, 0 \leq -q_{ii} < \infty$,
2. pour tous $i \neq j, q_{ij} \geq 0$,
3. pour tout $i \in I, \sum_{j \in I} q_{ij} = 0$.

On notera $q_i = q(i) = -q_{ii}$ pour tout $i \in I$. à un générateur Q , on associe la matrice de saut $\Pi = (\pi_{ij})_{i,j \in I}$ donnée par

$$\pi_{ii} = \begin{cases} 0 & \text{si } q \neq 0 \\ 1 & \text{si } q = 0 \end{cases}$$
$$\pi_{ij} = \begin{cases} q_{ij}/q_i & \text{si } j \neq i \text{ et } q \neq 0 \\ 1 & \text{si } q = 0 \end{cases}$$

Remarque 2.4

La matrice Π est stochastique : ces coefficients sont positifs et leur somme sur chaque ligne vaut 1 . Une chaîne de Markov à temps continu $(X_t)_{t \geq 0}$ est déterminée par une mesure λ sur I (identifiée à un vecteur ligne) et un générateur Q . La mesure λ détermine la distribution initiale (c'est la loi de X_0) et la matrice Q détermine la dynamique de la chaîne. On peut décrire de deux manières équivalentes la dynamique de $(X_t)_{t \geq 0}$. La première description se fait par l'intermédiaire de la chaîne incluse et des temps d'attente.

Définition 2.19

Le processus $(X_t)_{t \geq 0}$ est une chaîne de Markov à temps continu si, pour tout $n \in \mathbb{N}$, tous $0 \leq t_0 \leq t_1 \leq \dots \leq t_{n+1}$ et tous états $i_0, \dots, i_{n+1}$

$$P \left(X_{t_{n+1}=i_{n+1}} | X_{t_0=i_0}, \dots, X_{t_n=i_n} \right) = p_{i_n i_{n+1}}(t_{n+1} - t_n) .$$

2.9 Propriété de chaîne de Markov

On vérifie facilement que, pour toute chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ de probabilité de transition P et tout couple d'états (x, y) tel que $P(X_0 = x) > 0$, on a $P(x, y) = P(X_1 = y \mid X_0 = x)$. Plus généralement :

$$P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x) = P(x, y)$$

pour toute suite $(x_0, x_1, \dots, x_{n-1})$ d'états telle que

$$P(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x) > 0$$

On voit donc que la loi conditionnelle de X_{n+1} sachant $(X_0, \dots, X_n)$ ne dépend que de X_n :

$$\begin{aligned} P(x, y) &= P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x) \\ &= P(X_{n+1} = y \mid X_n = x) \\ &= P(X_1 = y \mid X_0 = x) \end{aligned}$$

C'est précisément le caractère markovien dont nous avons parlé en introduction. Le but de cette section est d'étendre cette propriété à des variables aléatoires autres que X_{n+1} et des temps d'arrêt non nécessairement constants.

Chaîne de markov forte

Définition 2.20

Soit $\{X_n; n \geq 0\}$ une chaîne de Markov. Pour simplifier les notations nous allons noter P_x la probabilité conditionnelle sachant $X_0 = x$, c'est-à-dire $P_x(A) = P(A \mid X_0 = x)$

$$F_n = \left\{ \{w; (X_0(w); \dots; X_n(w)) \in B_n\}; B_n \in \mathcal{P}(E^{n+1}) \right\}$$

Définition 2.21

Soit T une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$. On dira que T est un temps d'arrêt si pour tout $n \in \mathbb{N}$, $\{t = n\} \in \mathcal{F}_n$.

Théorème 2.3

Soit $\{X_n; n \geq 0\}$ une chaîne de Markov de loi initiale μ et de matrice de transition P et T un temps d'arrêt. Alors pour tout $A \in \mathcal{F}_T$ et pour tout $m \geq 1, x_1, \dots, x_m \in E$,

$$P(A \cap \{X_{T+1} = x_1; \dots; X_{T+m} = x_m\} \mid X_T = x; T < \infty) = P(A \mid X_T = x; T < \infty) P_x(X_1 = x_1;$$

ou $\mathcal{F}_T = \sigma(X_0; \dots; X_T)$.

Chaînes de Markov faibles

Notation

Dans toute la suite on notera F_n la tribu engendrée par les v.a. $X_0; \dots; X_n$, autrement dit $F_n = \sigma(X_0; \dots; X_n)$. On notera également P_x la probabilité conditionnelle sachant $\{X_0 = x\}$, $x \in E$.

Théorème de chaîne de Markov faible

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov. Alors pour tous $n \in \mathbb{N}, x \in E$ et conditionnellement à $\{X_n = x\}$, $(X_{n+p})_{p \in \mathbb{N}}$ est une chaîne de Markov de loi initiale δ_x . De plus elle est indépendante de F_n ; c'est-à-dire, pour tout $A \in F_n$, pour tous $m > 0, x_0; x_1; \dots; x_m \in E$

$$\begin{aligned} & P(A \cap \{X_{n+1} = x_1; \dots; X_{n+m} = x_m\} \mid X_n = x) \\ &= P(X_{n+1} = x_1; \dots; X_{n+m} = x_m \mid X_n = x) P(A \mid X_n = x) \\ &= P_x(X_1 = x_1; \dots; X_m = x_m) P(A \mid X_n = x) : \end{aligned}$$

Avant de faire la démonstration, nous allons établir 2 lemmes.

Lemme 2.2

Soit $(\omega; F; P)$ un espace probabilisé et soient $A; B; C \in F$ tels que $P(B \cap C) > 0$:
Alors

$$P(A|B|C) = P(A \mid B \cap C)$$

Démonstration : Soit $Q(\cdot) := P(\cdot \mid C)$. Q est une probabilité. Par définition de la probabilité conditionnelle, on a

$$\begin{aligned} P(A|B|C) &= Q(A \mid B) = \frac{Q(A \cap B)}{Q(B)} = \frac{P(A \cap B \mid C)}{P(B \mid C)} \\ &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= P(A \mid B \cap C). \end{aligned}$$

■

Lemme 2.3

Soit $(X_n)_{n \geq 0}$ une chaîne de Markov. Soit ≥ 0 et $x_0; x_n; x_{n+1} \in E$ tels que

$$P(\{X_n = x_n\} \cap \{X_0 = x_0\}) > 0.$$

Alors

$$P(X_{n+1} = x_{n+1} \mid X_n = x_n; X_0 = x_0) = P(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

[3]

Application de Chaînes de Markove

3.1 Le processus de Galton-Watson

Le processus de branchement discret ou processus de Galton-Watson est une chaîne de Markov homogène sur $\mathbb{N}$ qui décrit l'évolution de la taille des générations successives d'une population où chaque individu donne naissance à un nombre aléatoire d'individus. Dans ce modèle, on suppose que les variables aléatoires décrivant le nombre de descendants de chaque individu sont indépendantes et de même loi.

Définition et exemples

Définition 3.1

Soit μ une loi de probabilité sur $\mathbb{N}$ et $\{X_{n,i}, i \in \mathbb{N}^*, n \in \mathbb{N}\}$ une famille de variables aléatoires indépendantes de loi μ . Soit U une variable aléatoire sur $\mathbb{N}^*$ indépendante de la famille $\{X_{n,i}, i \in \mathbb{N}, n \in \mathbb{N}\}$. La suite de variables aléatoires $(Z_n)_n$ définie par :

$$Z_0 = U \text{ et pour } n \in \mathbb{N}, \quad Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} X_{n,i} & \text{si } Z_n \geq 1 \\ 0 & \text{si } Z_n = 0 \end{cases}$$

est appelée un processus de Galton-Watson de loi de reproduction μ .

Exemple 3.1

Division cellulaire : on considère une population de cellules toutes identiques (génération 0). Si on ne tient pas compte de l'apparition possible de mutations, chaque cellule de cette population donne naissance à deux cellules identiques avec probabilité p et meurt sans se diviser avec probabilité $1 - p$ ($p \in [0, 1]$). Les cellules issues de la division des cellules de la génération $n - 1$ constituent les cellules de la génération n . La loi de reproduction μ vérifie $\mu(0) = 1 - p$ et $\mu(2) = p$.

Exemple 3.2

Survivance des noms de familles : le processus de branchement a été introduit pour la première fois par Francis Galton en 1873 pour étudier la survivance des noms de famille en Grande-Bretagne. Comme les noms de famille sont transmis par le père,

$\mu(k)$ représente dans ce modèle, la probabilité qu'un homme ait k fils.

Propriétés

Dans la suite, $(Z_n)_n$ désigne un processus de Galton-Watson de loi de reproduction μ .

Proposition 3.1

La suite $(Z_n)_n$ est une chaîne de Markov homogène sur $\mathbb{N}$ de matrice de transition Q définie par $Q(i, j) = P\left(\sum_{k=1}^i X_{0,k} = j\right)$ pour tout $i, j \in \mathbb{N}$

- Si $\mu(1) \neq 1$, l'état 0 est un état absorbant et tous les autres états sont transitoires.
- Si $\mu(1) = 1$, tous les états sont absorbants.

Démonstration : Soit $n \in \mathbb{N}^*$ et $k_0, \dots, k_n \in \mathbb{N}$. D'après la définition de Z_n et en utilisant la convention que $\sum_{i=1}^0 a_i = 0$:

$$P(Z_0 = k_0, \dots, Z_n = k_n) = P\left(U = k_0, \sum_{i=1}^{k_0} X_{0,i} = k_1, \dots, \sum_{i=1}^{k_{n-1}} X_{n-1,i} = k_n\right).$$

Par hypothèse, $U, \{X_{0,i}, i \in \mathbb{N}\}, \dots, \{X_{n-1,i}, i \in \mathbb{N}\}$ sont toutes indépendantes. Donc,

$$P(U = k_0, \dots, Z_n = k_n) = P(U = k_0) P\left(\sum_{i=1}^{k_0} X_{0,i} = k_1\right) \dots P\left(\sum_{i=1}^{k_{n-1}} X_{n-1,i} = k_n\right).$$

Enfin, comme les variables aléatoires $X_{n,i}$ sont toutes de même loi μ ,

$$P(U = k_0, \dots, Z_n = k_n) = P(U = k_0) Q(k_0, k_1) \dots Q(k_{n-1}, k_n)$$

en posant $Q(k, \ell) = P\left(\sum_{i=1}^k X_{0,i} = \ell\right)$ pour tout $k, \ell \in \mathbb{N}$. Donc, $Q(0, j) = \mathbb{1}_{j=0}$ et $Q(i, j) = P(X_{0,1} + \dots + X_{0,i} = j)$ si $i \in \mathbb{N}^*$.

Nature des états : $\mu(1) = 1$ signifie que chaque individu donne naissance à un unique individu, la taille de la population reste fixe : tous les états sont absorbants. On suppose maintenant que $\mu(1) < 1$.

— Si $\mu(0) > 0$, alors pour tout $i \in \mathbb{N}^*$, $Q(i, 0) = \mu(0)^i > 0$ (c'est la probabilité que les i individus de la population n'aient pas de descendants). Comme 0 est un état absorbant, $P(\text{pour tout } n \in \mathbb{N}^*, Z_n \neq i \mid Z_0 = i) \geq Q(i, 0) > 0$. Donc, pour tout $i \in \mathbb{N}^*$, i est transitoire.

— Supposons que $\mu(0) = 0$. Cela signifie que tous les individus donnent naissance à au moins un individu avec probabilité 1. Comme $\mu(1) < 1$, chaque individu a une probabilité strictement positive d'avoir au moins 2 descendants. Donc, partant d'une population avec $i > 0$ individus, il y a une probabilité strictement positive que les

génération suivantes aient toujours au moins $i + 1$ individus : précisément, il existe $l \geq 2$ tel que $\mu(l) > 0$. Donc, $Q(i, il) > 0$ et comme $Q(i, j) = 0$ si $j < i$,

$$P(Z_n \neq i, \text{ pour tout } n > 0 \mid Z_0 = i) \geq P(X_{0,1} = l, \dots, X_{0,i} = l) = \mu(l)^i > 0.$$

Cela montre que i est transitoire. ■

Proposition 3.2 (*Comportement asymptotique*)

Notons Ext l'événement "la population s'éteint au bout d'un nombre fini de générations" :

$$Ext = \{ \text{Il existe } n \in \mathbb{N}, Z_n = 0 \} = \{T < +\infty\} \text{ où } T = \inf (n > 0, Z_n = 0).$$

Posons $q = P(Ext)$ la probabilité que la population s'éteigne.

- (i) La suite $(P(Z_n = 0))_n$ est croissante et converge vers q .
- (ii) Si $\mu(1) \neq 1$ alors avec probabilité 1, la population s'éteint ou sa taille tend vers $+\infty$ et donc

$$q + P\left(\left\{Z_k \xrightarrow[k \rightarrow +\infty]{} \infty\right\}\right) = 1.$$

Démonstration :

- (i) Par construction $Z_n = 0$ implique que $Z_{n+1} = 0$. Donc $(P(Z_n = 0))_n$ est une suite croissante. D'autre part, $q = P(T < +\infty) = \lim_{n \rightarrow +\infty} P(T \leq n)$. Or, pour tout $n \in \mathbb{N}^*$, $\{T \leq n\} = \{Z_n = 0\}$.

Donc, $(P(Z_n = 0))_n$ converge vers q

- (ii) Soit $N \in \mathbb{N}^*$. Comme tous les états $1, \dots, N - 1$ sont transitoires, $P(Z_n \in]0, N[\text{ pour une infinité de } n) \leq \sum_{k=1}^{N-1} P(Z_n \text{ passe une infinité de fois en } k) = 0$.

Donc, l'événement "il n'y a pas extinction de la population et il existe un entier $N > 0$ tel que la taille d'une infinité de générations soit inférieure à N " est de probabilité nulle : $P(\text{ Il existe } N \in \mathbb{N}^*, Z_n \in]0, N[\text{ pour une infinité de } n) \leq \sum_{N=1}^{+\infty} P(Z_n \in]0, N[\text{ pour une infinité de } n) = 0$.

L'événement contraire est "au bout d'un nombre fini de générations, la population s'éteint ou pour tout entier N , on peut trouver une génération à partir de laquelle la taille de la population dépasse toujours N ". Il arrive avec probabilité 1. Donc, $P(Ext \cup \left\{Z_n \xrightarrow[n \rightarrow +\infty]{} +\infty\right\}) = 1$.

Comme les deux événements "extinction de la population" et "explosion de la taille de la population" sont incompatibles, on peut écrire le résultat différemment

$$q + P\left(Z_n \xrightarrow[n \rightarrow +\infty]{} \infty\right) = 1.$$



Evolution en moyenne de la taille de la population

On suppose dans ce paragraphe que la loi de reproduction μ a une espérance finie notée m $\left(m = \sum_{k=0}^{+\infty} k\mu(k)\right)$ et que la taille initiale de la population Z_0 est d'espérance finie.

Proposition 3.3

Pour tout $n \in \mathbb{N}^*$, la taille moyenne d'une population à la n -ième génération est

$$E(Z_n) = m^n E(Z_0)$$

Démonstration : Soit $k \in \mathbb{N}$. Pour calculer $E(Z_n)$, on va d'abord calculer l'espérance conditionnelle $E(Z_n | Z_{n-1} = k)$ pour tout $k \in \mathbb{N}$. S'il n'y a plus d'individu dans la population à la $(n-1)$ -ième génération alors c'est encore le cas à la n -ième génération.

Donc, $E(Z_n | Z_{n-1} = 0) = 0$. Si $k > 0$, $E(Z_n | Z_{n-1} = k) = E\left(\sum_{l=1}^k X_{n-1,l} | Z_{n-1} = k\right)$.

D'après la relation (10), Z_{n-1} dépend uniquement des variables aléatoires $X_{k,i}$ pour $k < n-1$ et $i \in \mathbb{N}^*$. Donc, les variables aléatoires $X_{n-1,l}$ sont indépendantes de Z_{n-1} .

D'autre part, elles sont de loi μ . Donc, $E(Z_n | Z_{n-1} = k) = mk$. Calculons $E(Z_n)$:

$$E(Z_n) = \sum_{k=0}^{+\infty} E(Z_n | Z_{n-1} = k) P(Z_{n-1} = k) = \sum_{k=1}^{+\infty} mk P(Z_{n-1} = k) = m E(Z_{n-1})$$

Par conséquent, $E(Z_n) = m^n E(Z_0)$ pour tout $n \in \mathbb{N}$. ■

Définition 3.2

Le processus de Galton-Watson $(Z_n)_n$ est dit sous-critique si $m < 1$, critique si $m = 1$ et sur-critique si $m > 1$.

N.B. Si $m < 1$ et si $E(Z_0) < +\infty$, $(E(Z_n))_n$ tend vers 0 et donc la population s'éteint en un nombre fini de générations avec probabilité 1. En effet, pour tout n , $P(Z_n \geq 1) \leq E(Z_n)$, d'après l'inégalité de Markov.

Donc, $(P(Z_n = 0))_n$ tend vers 1. Comme la probabilité que la population s'éteigne en un nombre fini de générations est minorée par $P(Z_n = 0)$ pour tout n , elle vaut 1.

Probabilité d'extinction de la population issue d'un seul ancêtre

On suppose dans ce paragraphe qu'initialement la population est formée d'une seule personne : $Z_0 = 1$. On cherche à calculer la probabilité que la population s'éteigne : $q = P(\text{Ext})$.

Théorème 3.1

Notons f l'application définie sur $[-1, 1]$ par $f(x) = \sum_{k=0}^{+\infty} x^k \mu(k)$

- Pour tout $n \in \mathbb{N}^*$, $P(Z_n = 0) = f(P(Z_{n-1} = 0))$.
- La probabilité d'extinction q est la plus petite solution positive de l'équation $x = f(x)$.

Démonstration :

- Notons $\mathcal{S}$ le support de la loi de Z_1 c'est-à-dire l'ensemble des entiers k tels que $P(Z_1 = k) > 0$. Alors,

$$P(Z_n = 0) = \sum_{k \in \mathcal{S}} P(Z_n = 0 \mid Z_1 = k) P(Z_1 = k).$$

Comme $Z_0 = 1$, la loi de Z_1 est μ . Supposons que $Z_1 = k$ avec $k \in \mathbb{N}^*$, c'est-à-dire que la population à la génération 1 est constituée de k individus $I_1, \dots, I_k$. Notons $V_n^{(i)}$ le nombre d'individus de la génération n qui sont des descendants de I_i . Alors, la loi conditionnelle de Z_n sachant que $Z_1 = k$ est la même que la loi de $\sum_{i=1}^k V_n^{(i)}$. D'autre part, les variables aléatoires $V_n^{(1)}, \dots, V_n^{(k)}$ sont indépendantes et de même loi que Z_{n-1} . Donc,

$$P(Z_n = 0 \mid Z_1 = k) = P\left(\sum_{i=1}^k V_n^{(i)} = 0\right) = P\left(\{V_n^{(1)} = 0\} \cap \dots \cap \{V_n^{(k)} = 0\}\right) = P(Z_{n-1} = 0)^k.$$

Cela montre que $P(Z_n = 0) = f(P(Z_{n-1} = 0))$

- On fait tendre n vers $+\infty$ dans les deux membres de la relation de récurrence. En utilisant que la suite $(P(Z_n = 0))_n$ tend vers q et le théorème de convergence dominée, on obtient que $\sum_{k=0}^{+\infty} P(Z_{n-1} = 0)^k \mu(k)$ converge vers $\sum_{k=0}^{+\infty} q^k \mu(k)$. Cela montre que $q = f(q)$. Il reste à montrer que q est la plus petite solution positive de cette équation. Soit e un réel positif vérifiant $e = f(e)$, montrons que pour tout $n \in \mathbb{N}^*$, $P(Z_n = 0) \leq e$.
On a $P(Z_1 = 0) = \mu(0) = f(0) \leq f(e) = e$ car f est croissante sur $[0, 1]$. Supposons que $P(Z_n = 0) \leq e$. Alors, $P(Z_{n+1} = 0) = f(P(Z_n = 0)) \leq f(e) = e$. Donc, $P(Z_n = 0) \leq e$ pour tout $n \in \mathbb{N}^*$. Comme $(P(Z_n = 0))_n$ tend vers q , on en déduit que $q \leq e$. ■

Définition 3.3

La fonction $f : s \mapsto \sum_{k=0}^{+\infty} s^k \mu(k)$ est appelée la fonction génératrice de la loi de reproduction μ (si X est une variable aléatoire de loi μ alors $f(s) = E(s^X)$ pour $s \in [-1, 1]$, f

est aussi appelée la fonction génératrice de X).

Remarquons que 1 est toujours solution de l'équation $x = f(x)$. Une étude de la fonction f permet d'obtenir le résultat suivant :

Corollaire 3.1

Notons m l'espérance de la loi de reproduction ($m \in \mathbb{R}_+ \cup \{+\infty\}$).

- Si $\mu(0) = 0$, alors $q = 0$.
- Si $\mu(0) > 0$ et $m \leq 1$, alors $q = 1$.
- Si $\mu(0) > 0$ et $m > 1$, alors $q \in]0, 1[$ et c'est l'unique solution dans l'intervalle $[0, 1[$ de l'équation $f(s) = s$.

Démonstration :

- Si $\mu(0) = 0$, chaque individu de la population donne naissance à au moins un individu avec probabilité 1. Donc, pour tout $n \in \mathbb{N}^*$, $P(Z_n = 0 \mid Z_0 = 1) = 0$, cela implique que la population a une probabilité nulle de s'éteindre.
- Si $\mu(0) > 0$ et si $\mu(0) + \mu(1) = 1$, alors $f(s) = \mu(0) + (1 - \mu(0))s$ et l'unique solution de $f(s) = s$ est 1. Remarquons que dans ce cas $m < 1$.
- Supposons maintenant que $\mu(0) > 0$ et que $\mu(0) + \mu(1) < 1$. Posons $h(s) = f(s) - s$. On a $h(0) = \mu(0) \in]0, 1[$ et $h(1) = 0$. On va utiliser les propriétés suivantes de f (communes à toutes les fonctions génératrices de loi de probabilité sur $\mathbb{N}$) :
- la fonction f est au moins dérivable deux fois sur $] - 1, 1[$: pour tout $s \in] - 1, 1[$,

$$f'(s) = \sum_{k=1}^{+\infty} k s^{k-1} \mu(k) f''(s) = \sum_{k=2}^{+\infty} k(k-1) s^{k-2} \mu(k).$$

- f' est donc croissante sur $[0, 1[$.
- $\lim_{s \uparrow 1} f'(s) = m$ Pour $s \in [0, 1[$, h est donc deux fois dérivable, $h'(s) = f'(s) - 1$ et $h''(s) = f''(s)$. Comme il existe $k \geq 2$ tel que $\mu(k) > 0$, h'' est strictement positive sur $[0, 1[$ et donc la fonction h' est strictement croissante sur l'intervalle $[0, 1[$. On a $h'(0) = \mu(1) - 1 < 0$ et $h'(s)$ tend vers $m - 1$ lorsque s tend vers 1. Le comportement de h est différent suivant que $m \leq 1$ ou $m > 1$ (faire un tableau de variations pour h).
- Si $m \leq 1$ alors h' est strictement négative sur $[0, 1[$ et donc h est strictement décroissante sur l'intervalle $] 0, 1[$. Par conséquent, la seule solution de $f(s) = s$ dans l'intervalle $[0, 1]$ est 1.
- Si $m > 1$ alors il existe un unique $a \in]0, 1[$ tel que $h'(a) = 0$. Cela montre que $h(s) = 0$ a une unique solution dans l'intervalle $[0, 1[$.

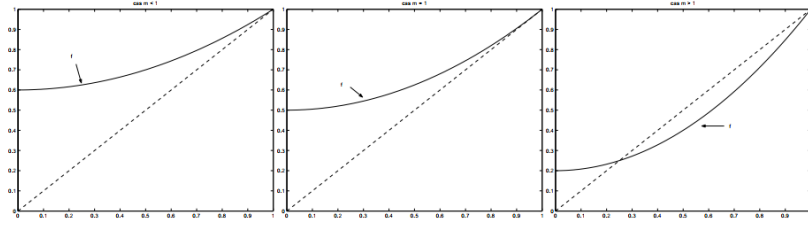


FIGURE 3.1 – Position de la courbe de f par rapport à la diagonale suivant que $m < 1$, $m = 1$ ou $m > 1$ lorsque $\mu(0) > 0$ et $\mu(0) + \mu(1) < 1$

■

Exemple 3.3

Division cellulaire : $\mu(0) = 1 - p$ et $\mu(2) = p$. La fonction génératrice de μ est la fonction f définie par $f(s) = 1 - p + ps^2$. La probabilité d'extinction est égale à $(1 - p)/p$ si $p > 1/2$ et à 1 si $p \leq 1/2$.

Exemple 3.4

Loi de reproduction géométrique : $\mu(k) = p(1 - p)^k$ pour $k \in \mathbb{N}$. La fonction génératrice de μ est la fonction f définie sur $[-1, 1]$ par

$$f(s) = \sum_{k=0}^{+\infty} s^k \mu(k) = p \sum_{k=0}^{+\infty} (s(1 - p))^k = \frac{p}{1 - s(1 - p)}.$$

La probabilité d'extinction est égale à $p/(1 - p)$ si $p < 1/2$ et à 1 si $p \geq 1/2$.

Probabilité d'extinction d'un processus de Galton-Watson issu de k ancêtres

Soit $(Z_n)_n$ un processus de Galton-Watson de loi de reproduction μ issu de $k \geq 2$ ancêtres ($Z_0 = k$). Notons $I_1, \dots, I_k$ les individus de la population initiale. Pour $i \in \{1, \dots, k\}$, posons $Z_0^{(i)} = 1$ et pour $n \geq 1$, posons $Z_n^{(i)}$ le nombre de descendants de la génération n issus de l'individu I_i . Les suites $(Z_n^{(i)})_{n \geq 1}$ sont des processus de Galton-

Watson indépendants issus d'un ancêtre et de loi de reproduction μ et $Z_n = \sum_{i=1}^k Z_n^{(i)}$ pour tout $n \in \mathbb{N}$. En particulier, le processus $(Z_n)_n$ s'éteint signifie que les populations issues des individus $I_1, \dots, I_k$ s'éteignent toutes.

En utilisant l'indépendance des suites $(Z_n^{(i)})_n$, on obtient :

Corollaire 3.2

La probabilité que le processus de Galton-Watson $(Z_n)_n$ issu de k ancêtres s'éteigne est q^k où q est la plus petite solution positive de l'équation $f(s) = s$, f étant la fonction génératrice de la loi de reproduction.

Démonstration : $P(\text{il existe } n \in \mathbb{N}^*, Z_n = 0) = P(\text{pour tout } i \in \{1, \dots, k\}, \text{il existe } n_i \in \mathbb{N}, Z_n^{(i)} = 0) = \prod_{i=1}^k P(\text{il existe } n \in \mathbb{N}^*, Z_n^{(i)} = 0) = q^k.$ ■

Complément : comportement asymptotique du processus de Galton-Watson dans le cas sur-critique

Soit $(Z_n)_n$ un processus de Galton-Watson, issu d'un ancêtre ($Z_0 = 1$), de loi de reproduction μ . On suppose que la loi de reproduction μ a une espérance finie $m > 1$ (cas sur-critique) et qu'elle admet une variance σ^2 finie. Le théorème suivant décrit le comportement asymptotique de la suite $(W_n)_n$ définie par $W_n = \frac{Z_n}{m^n}$ pour tout $n \in \mathbb{N}$

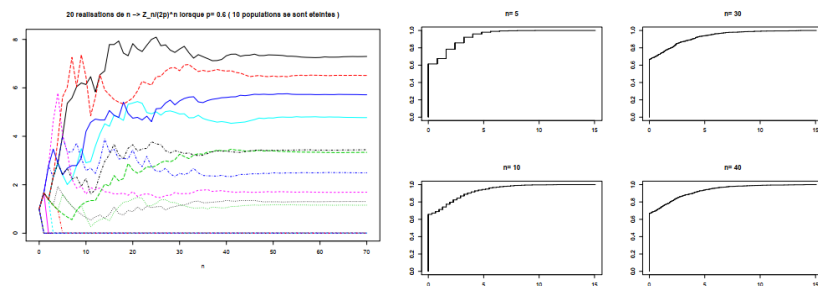
Théorème 3.2

La suite $(W_n)_n$ converge avec probabilité 1 vers une variable aléatoire W . De plus,

- (i) $E((W_n - W)^2) \xrightarrow{n \rightarrow +\infty} 0.$
- (ii) $E(W) = 1$ et $\text{Var}(W) = \frac{\sigma^2}{m^2 - m}.$
- (iii) $P(W = 0) = q$ où q est la probabilité d'extinction du processus de Galton-Watson $(Z_n)_n$.

Exemple 3.5

Etude par simulations de l'évolution de la taille d'une population de bactéries obtenues par divisions cellulaires à partir d'une seule bactérie. On suppose que chaque bactérie a une probabilité $p = 0.6$ de se diviser. La variable aléatoire Z_n est le nombre de cellules de la génération n . Ici $m = 2p$ et la probabilité d'extinction de la population est $q = 2/3$.



Sur la figure de gauche, on observe bien que chaque réalisation de $n \rightarrow \frac{\ln(Z_n)}{(2p)^n}$ semble se stabiliser lorsque n est grand vers une valeur qui dépend de la réalisation. Les quatre figures de droite montrent la fonction de répartition empirique de 1000 réalisations de $W_n = \frac{Z_n}{(2p)^n}$ pour quatre valeurs de n ; on observe bien que la fonction de répartition empirique se stabilise lorsque n augmente; la proportion de réalisations nulles de W_{40} est 0.668, la moyenne empirique des 1000 réalisations de W_{40} est de 1.029 et sa variance empirique est de 4.325. La courbe "limite" doit être le graphe de la fonction de répartition de la variable aléatoire W introduite dans le théorème 3.2 (ici $q = 2/3$, $E(W) = 1$ et $\text{Var}(W) = 4$).

Bibliographie

- [1] ANDRÉ MONIM. *Processus stochastique estimation filtrag*. université de toulouse. 09/2009.
- [2] AMÉLIE GUILBAUT. *Chaînes de markov cachées*. université de lille 1. 11 mai 2018.
- [3] BECHKIT ABDERAHMANE. *Processus de sauts de markov application a la $M = E^r = 1$* . université de m'hamed bougarra .28/06/2017.
- [4] BTS DOMOTIQUE. *Fiabilité*. université de mosd sil 2008-2010.
- [5] CLAUDIE HASSENFORDER ISMAG MASTER. *Processus stochastiques modélisation*. université de toulous.
- [6] DJALIL CHAFAÏ. *Quelques aspects des chaînes de Markov*. université telemcen .18 janvier 2007.
- [7] JEAN-YVES DAUXIOS. *Recherche de mathématiques probabilités*. université de franche-comté. année scolaire 2008-2009.
- [8] JEAN-CHRISTOPHE BRETON. *Processus stochastiques*. université de rennes 1. octobre 2018.
- [9] NILS BERGLUND. *Processus aléatoires et applications*. université d'orléans. janvier 2014.
- [10] OMAR BELABBCI. *La penalisation des trajectoire du mouvement browien*. universite aboubeker belkaid - tLemcen. 2008 .