DEA Environnement Marin : processus stochastiques

Avner Bar-Hen

Université Aix-Marseille II

2001-2002

Table des matières

1 Rappels de calcul matriciel et de probabilité (cas discret) 3

1 Matrices . . . 3
1.1 Définition . . . 3
1.2 Opérations sur les matrices . . . 4
1.3 Matrice identité . . . 4
1.4 Matrice inverse . . . 5
1.5 Matrice diagonalisable . . . 5
1.6 Puissance et exponentielle d'une matrice . . . 5
2 Probabilités . . . 5
2.1 Axiomes de probabilité . . . 5
2.2 Probabilités conditionnelles et événements indépendants . . . 6
2.3 Variables aléatoires discrètes . . . 7
3 Quelques lois discrètes . . . 8
3.1 Quelques séries à connaître . . . 8
3.2 Loi binomiale . . . 9
3.3 Loi de Poisson (loi des événements rares) . . . 9
3.4 Loi géométrique ou loi de Pascal . . . 9
3.5 Loi géométrique généralisée ou loi binomiale négative . . . 10
3.6 Loi hypergéométrique . . . 10
4 Exercices . . . 11
4.1 Calcul . . . 11
4.2 Puissance . . . 11
4.3 Voisins . . . 11
4.4 Parapluie . . . 12
4.5 Moutons . . . 12
4.6 Dur métier . . . 12

2 Couple de variables aléatoires, loi marginale, loi conditionnelle 13

1 Couple de variables aléatoires . . . 13
1.1 Espérance du couple (X, Y) . . . 13
1.2 Matrice de variance-covariance du couple (X, Y) . . . 13
2 Loi marginale . . . 14
3 Loi conditionnelle . . . 14
4 Espérance conditionnelle . . . 14

5 Exercices . . . 16
5.1 Lois marginales et conditionnelles . . . 16
5.2 Binomiale et Poisson . . . 16
5.3 Famille nombreuse . . . 16

3 Chaînes de Markov 17
1 Définitions . . . 17
2 Propriété fondamentale des chaînes de Markov homogènes . . . 19
3 Calcul de la loi de X_n . . . 19
4 Matrice de transition dans le temps . . . 20
5 Exercices . . . 22
5.1 Weather in the Land of Oz . . . 22
5.2 Propriétés des chaînes de Markov homogènes . . . 22
5.3 Chaîne à deux états . . . 22
5.4 Météo . . . 23

4 Convergence des chaînes de Markov homogènes 25
1 Probabilité invariante et convergence . . . 26
2 Classification des chaînes de Markov . . . 26
2.1 Exemple : marche aléatoire . . . 29
3 Résultat fondamental pour les chaînes à espace d'états fini . . . 30
4 Exercices . . . 31
4.1 Weather in the Land of Oz . . . 31
4.2 Chaîne produit . . . 31
4.3 Reproduction diploide . . . 31
4.4 Chaîne de Wright . . . 32
4.5 Chaîne d'Ehrenfest . . . 33

5 Processus de branchement (ou de ramification ou de Galton-Watson) 35
1 Description du proccessus . . . 35
1.1 Fonction génératrice . . . 36
1.2 Étude du processus de branchement (Y_n) . . . 36
2 Exercices . . . 37

Chapitre 1

Rappels de calcul matriciel et de probabilité (cas discret)

1 Matrices

1.1 *Définition*

Une *matrice* $p \times q$ est un tableau à p lignes et q colonnes. On note $A = (a_{ij})_{1 \leq i \leq p; 1 \leq j \leq q}$ la matrice dont a_{ij} est l'élément de la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne.

Exemple : $A = \begin{pmatrix} 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 1 & 0 \\ 1 & 2 & 7 & 0 & 8 \end{pmatrix}$ est une matrice 3×5 et $a_{1,5} = 3$.

Un vecteur de $\mathbb{R}^p$ est une matrice $p \times 1$. Dans ce cas on note $v = (v_i)_{1 \leq i \leq p}$.

Exemple : $v = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \\ 8 \end{pmatrix}$ est un vecteur de $\mathbb{R}^5$.

La *transposée* de la matrice $A = (a_{ij})_{1 \leq i \leq p; 1 \leq j \leq q}$ est une matrice $q \times p$ noté A^* et définie par : $A^* = (b_{ij})_{1 \leq i \leq q; 1 \leq j \leq p}$ avec $b_{ij} = a_{ji}$.

Exemple : $A^* = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 7 \\ 2 & 1 & 0 \\ 3 & 0 & 8 \end{pmatrix}$

Une *matrice carrée* est une matrice ayant le même nombre de lignes et de colonnes. Une matrice carrée $n \times n$ est dite d'*ordre* n. Une matrice est dite *symétrique* si elle est égale à sa transposée.

Exemple : $B = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \\ 1 & 2 & 7 & 0 \\ 2 & 1 & 0 & 3 \end{pmatrix}$ est une matrice symétrique.

Une matrice carrée $A = (a_{ij})_{1 \le i \le n; 1 \le j \le n}$ est dite *diagonale* si $a_{ij} = 0 \ \forall i \neq j$.

1.2 Opérations sur les matrices

Addition de 2 matrices $p \times q$

Soient $A = (a_{ij})_{1 \le i \le p; 1 \le j \le q}$ et $B = (b_{ij})_{1 \le i \le p; 1 \le j \le q}$ deux matrices $p \times q$.

La somme de A et de B, notée $A + B$, est la matrice $p \times q$ définie par : $A + B = (c_{ij})_{1 \le i \le p; 1 \le j \le q}$ avec $c_{ij} = a_{ij} + b_{ij}$.

Il est clair que l'addition des matrices est commutative : $A + B = B + A$ et associative : $(A + B) + C = A + (B + C)$

Multiplication d'une matrice $p \times q$ **et d'un scalaire** λ

Soit $A = (a_{ij})_{1 \le i \le p; 1 \le j \le q}$ une matrice $p \times q$ et soit un scalaire λ.

La multiplication de A et de λ, notée λA, est une matrice $p \times q$ définie par : $\lambda A = (c_{ij})_{1 \le i \le p; 1 \le j \le q}$ avec $c_{ij} = \lambda a_{ij}$

Multiplication d'une matrice $p \times q$ **et d'une matrice** $q \times r$

Soient $A = (a_{ij})_{1 \le i \le p; 1 \le j \le q}$ une matrice $p \times q$ et $B = (a_{ij})_{1 \le i \le q; 1 \le j \le r}$. La multiplication de A et de B, notée AB, est une matrice $p \times r$ définie par $AB = (c_{ij})_{1 \le i \le p; 1 \le j \le r}$ avec $c_{ij} = \sum_{k=1}^{q} a_{ik} b_{kj}$

Remarques :

- La multiplication de A et de B n'est possible que si le nombre de colonnes de A est égal au nombre de lignes de B.
- Si A et B sont deux matrices carrées alors les produits AB et BA ont un sens mais en général $AB \neq BA$. (La multiplication des matrices n'est pas commutative).
- La multiplication des matrices est associative : $(AB)C = A(BC)$ et est distributive par rapport à l'addition : $A(B + C) = (AB) + (AC)$.
- On a la propriété $(AB)^* = B^* A^*$

1.3 Matrice identité

La matrice appelée *identité d'ordre* n et notée I_n est la matrice carrée d'ordre n ayant des 1 sur la diagonale et des 0 partout ailleurs.

Exemple : $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est la matrice identité d'ordre 3.

La matrice I_n vérifie $AI_n = I_n A = A$ pour toute matrice d'ordre n (I_n est l'élément neutre pour la multiplication dans l'ensemble des matrices carrées d'ordre n).

1.4 Matrice inverse

Soit A une matrice carrée d'ordre n. On dit que A est inversible si il existe une matrice B carrée d'ordre n telle que $AB = BA = I_n$. On dit alors que B est l'inverse de A et on note $B = A^{-1}$.
Pour les personnes connaissant les déterminants :
une matrice A est inversible si et seulement si $\det(A) \neq 0$ et on a : $A^{-1} = \frac{B}{\det(A)}$ où B est la matrice des cofacteurs de A : $B = (b_{ij})_{i \leq i, j \leq n}$ avec $b_{ij} = (-1)^{i+j}\Delta_{ij}$ et Δ_{ij} est le déterminant de la matrice $(n-1) \times (n-1)$ égale à la matrice A à laquelle on a enlevé la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne.

1.5 Matrice diagonalisable

Soit A une matrice carrée d'ordre n. On dit que A est diagonalisable s'il existe une matrice inversible P telle que $A = PDP^{-1}$ où D est une matrice diagonale. Une matrice symétrique est diagonalisable.

1.6 Puissance et exponentielle d'une matrice

Soit A une matrice carrée d'ordre n. La puissance $m^{\text{ème}}$ (où $m \in \mathbb{N}^*$) de la matrice A, notée A^m, est la matrice carrée d'ordre n définie par :

$$A^m = \underbrace{AA \cdots A}_{m \text{ fois}}$$

On remarque que $A^m = A^{m-1}A = AA^{m-1}$. Par convention $A^0 = I_n$.
L'exponentielle de la matrice A, notée $\exp(A)$, est la matrice carrée d'ordre n définie par :

$$\exp(A) = \sum_{m=0}^{+\infty} \frac{A^m}{m!}$$

Dans la suite de ce cours nous aurons à calculer des puissances de matrice symétrique, c'est-à-dire calculer A^m où A est une matrice symétrique :

- Si l'on sait diagonaliser A, on a $A = PDP^{-1}$ où D est une matrice diagonale. Dans ce cas on a : $A^m = PD^mP^{-1}$
- sinon il faut mettre au point une procédure itérative : $A^0 = I_n$ et $A^m = AA^{m-1}$

2 Probabilités

2.1 Axiomes de probabilité

1. Pour tout événement aléatoire A, on a : $0 \leq \mathbb{P}(A) \leq 1$
2. $\mathbb{P}(\Omega) = 1$

3. Pour tous événements aléatoires A et B tels que $A \cap B = \emptyset$ on a : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Ce dernier point se généralise facilement : Soient $A_1, A_2, \ldots, A_n, \ldots$ des événements aléatoires 2 à 2 incompatibles ($A_i \cap A_j = \emptyset$ si $i \neq j$) alors $\mathbb{P}\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} \mathbb{P}(A_i)$. Cet axiome s'appelle la σ-additivité.

Conséquences des axiomes

1. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
2. $\mathbb{P}(\emptyset) = 0$
3. Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$
4. $\mathbb{P}(B_1 \cup \ldots \cup B_n) \leq \mathbb{P}(B_1) + \ldots \mathbb{P}(B_n)$
5. Propriété de la partition disjointe : soient $B_1, B_2, \ldots, B_n)$ tels que $B_i \cap B_j = \emptyset$ pour $i \neq j$ et $B_1 \cap \ldots \cap B_n = \Omega$.
   On dit que $(B_1, B_2, \ldots, B_n)$ forme une partition disjointe de Ω. Une telle partition vérifie la propriété suivante :

   $$\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \ldots + \mathbb{P}(A \cap B_n)$$

   Cas particulier :
   B et $\overline{B}$ forment une partition disjointe de Ω donc $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$

2.2 *Probabilités conditionnelles et événements indépendants*

Soit B un événement aléatoire tel que $\mathbb{P}(B) \neq \emptyset$. On appelle *probabilité de A sachant B* et on note $\mathbb{P}(A|B)$ la quantité $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. On montre que $\mathbb{P}(A|B)$ est une probabilité.
Deux événements A et B sont dits *indépendants* si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
Remarquons que si $\mathbb{P}(B) \neq 0$ alors (A et B indépendants) $\Leftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$ (notion intuitive : la connaissance de B n'influe pas sur la réalisation de A).

Formule des probabilités composées

$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$

Formule de Bayes

Soit $(B_1, B_2, \ldots, B_n)$ une partition disjointe de Ω.
On a $\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \ldots + \mathbb{P}(A \cap B_n)$.
En particulier $\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B})$.

2.3 Variables aléatoires discrètes

On appelle variable aléatoire toute application de Ω dans $\mathbb{R}$.
Une *variable aléatoire discrète* est une variable aléatoire qui ne prend qu'un nombre dénombrable de valeurs. La *loi de probabilité* d'une variable aléatoire discrète X est donnée par :

$$\{\mathbb{P}(X = x) \ ; \ x \text{ valeur possible pour } X\}$$

Exemple : jet de deux dés : $\Omega = \{(x, y) \ ; \ x, y \in \{1, 2, \ldots, 6\}\}$. On note X la variable aléatoire égale à la somme de ces deux dés.

$$X : \begin{cases} \Omega \mapsto \{2, 3, \ldots, 12\} \\ (x, y) \mapsto x + y \end{cases}$$

X est une variable aléatoire discrète sa loi de probabilité est donnée par les 11 probabilités :

$\mathbb{P}(X = 2)$	$\mathbb{P}(X = 3)$	$\mathbb{P}(X = 4)$	$\mathbb{P}(X = 5)$	$\mathbb{P}(X = 6)$	$\mathbb{P}(X = 7)$	$\mathbb{P}(X = 8)$	$\mathbb{P}(X = 9)$	$\mathbb{P}(X = 10)$	$\mathbb{P}(X = 11)$	$\mathbb{P}(X = 12)$
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Espérance et variance d'une variable aléatoire discrète

On appelle *espérance* de X et on note $\mathbb{E}(X)$ la quantité

$$\mathbb{E}(X) = \sum x\mathbb{P}(X = x)$$

La sommation est sur tous les x valeurs possibles pour X.
L'espérance d'une variable aléatoire s'interprète comme sa valeur moyenne.
Soit $\Phi \ : \mathbb{R} \mapsto \mathbb{R}$. $\Phi(x)$ est une variable aléatoire discrète et on a :

$$\mathbb{E}(\Phi(X)) = \sum \Phi(x)\mathbb{P}(X = x)$$

La sommation est sur tous les x valeurs possibles pour X.
On n'a donc pas besoin de calculer la loi de probabilité de $\Phi(x)$ pour calculer $\mathbb{E}(\Phi(X))$.
On appelle *variance* de X et on note $\mathbb{V}(X)$ la quantité :

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$$

On appelle *écart-type* de X et on note $\sigma(X)$ la quantité $\sigma(X) = \sqrt{\mathbb{V}(X)}$. L'écart-type s'interprète comme la dispersion moyenne de X autour de sa moyenne.

Propriétés de l'espérance et de la variance

1. Soit X une variable aléatoire et soient $a, b \in \mathbb{R}$ alors $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.
2. Soit X une variable aléatoire ne prenant que des valeurs positives ($X \geq 0$) alors $\mathbb{E}(X) \geq 0$.
3. Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires alors $\mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}(X_i)$
4. $\mathbb{V}(X) \geq 0$. De plus $\mathbb{V}(X) = 0 \Rightarrow X$ non aléatoire.
5. Soit X une variable aléatoire et soient $a, b \in \mathbb{R}$ alors $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$

Indépendance de variables aléatoires discrètes

Soient X et Y deux variables aléatoires discrètes : $X \in \{x_1, x_2, \dots, x_k\}$ et $Y \in \{y_1, y_2, \dots, y_m\}$. On dit que X et Y sont *indépendantes* si et seulement si :

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j) \quad \forall i \in \{1, 2, \dots, k\}, \ \forall j \in \{1, 2, \dots, m\}$$

Soient $X_1, X_2, \dots, X_n$ n variables aléatoires discrètes. On dit que ces variables sont *indépendantes entre elles* si et seulement si :

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)\dots\mathbb{P}(X_n = x_n) \ \forall x_1, x_2, \dots, x_n$$

Propriétés des variables aléatoires indépendantes

- Soient $X_1, X_2, \dots, X_n$ des variables aléatoires **indépendantes** alors
$$\mathbb{E}(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} \mathbb{E}(X_i)$$
- Soient $X_1, X_2, \dots, X_n$ des variables aléatoires **indépendantes** alors
$$\mathbb{V}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathbb{V}(X_i)$$

3 Quelques lois discrètes

3.1 *Quelques séries à connaître*

Série exponentielle

$$\forall x \in \mathbb{R} \qquad \sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x$$

On rappelle que $k! = k \times (k-1) \times \dots \times 3 \times 2 \times 1$

Formule du binôme

$$\forall a, b \in \mathbb{R} \qquad \sum_{k=0}^{n} C_n^k a^k b^{n-k} = (a+b)^n$$

On rappelle que $C_n^k = \frac{n!}{k!(n-k)!}$

Séries géométriques

- $\forall x \in \mathbb{R} \ \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$ et $\forall x \in]-1, 1[\ \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$
- $\forall x \in \mathbb{R} \ \sum_{k=1}^{n} x^k = \frac{x-x^{n+1}}{1-x}$ et $\forall x \in]-1, 1[\ \sum_{k=1}^{+\infty} x^k = \frac{x}{1-x}$
- et en dérivant $\forall x \in]-1, 1[\ \sum_{k=0}^{+\infty} kx^{k-1} = \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$

3.2 Loi binomiale

On considère une expérience aléatoire ayant deux issues : A (succès) et $\overline{A}$ (échec) avec $\mathbb{P}(A) = p$ et $\mathbb{P}(\overline{A}) = q = 1 - p$.
On répète n fois cette expérience de manière identique et indépendante et on note : S_n le nombre de fois où l'on a observé A au cours des n expériences ; c'est-à-dire le nombre de succès au cours des n expériences.
On a $S_n \in \{0, 1, \dots, \}$ et :

$$\mathbb{P}(S_n = k) = C_n^k p^k (1-p)^{n-k}$$

On dit que S_n suit une loi binomiale de paramètres n et p et on note $S_n \sim \mathcal{B}(n, p)$.
On montre que $\mathbb{E}(S_n) = np$ et $\mathbb{V}(S_n) = np(1-p)$.
Remarquons que l'on peut écrire $S_n = X_1 + X_2 + \cdots + X_n$ où les X_i sont des variables indépendantes telles que :

$$X_i = \begin{cases} 1 \text{ si le résultat de la } i^{\text{ème}} \text{ expérience est } A \\ 0 \text{ si le résultat de la } i^{\text{ème}} \text{ expérience est } \overline{A} \end{cases}$$

Ces variables vérifient $\mathbb{P}(X_i = 1) = p$ et $\mathbb{P}(X_i = 0) = q = 1 - p$.

3.3 Loi de Poisson (loi des événements rares)

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ de loi de probabilité :

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad \forall k \in \mathbb{N}$$

où λ est un réel positif. On dit que X suit une loi de Poisson de paramètre λ et on note $X \sim \mathcal{P}(\lambda)$. On montre que $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$; Remarquons que lorsque $n \to +\infty$, $p \to 0$, $np \to \lambda$ (en pratique $n \geq 100$, $0.1 \leq np \leq 10$) alors la variable binomiale $S_n \sim \mathcal{B}(n, p)$ peut être approchée par la loi de Poisson $X \sim \mathcal{P}(\lambda)$. En pratique cela revient à considérer que $\mathbb{P}(S_n = k) \approx \mathbb{P}(X = k)$.

3.4 Loi géométrique ou loi de Pascal

On considère une expérience aléatoire à deux issues : A (succès) et $\overline{A}$ (échec) avec $\mathbb{P}(A) = p$ et $\mathbb{P}(\overline{A}) = q = 1 - p$. On répète cette expérience de manière identique et indépendantes et on note Y le nombre d'épreuves nécessaires pour obtenir le premier succès. On a :

$$\mathbb{P}(Y = k) = (1-p)^{k-1} p \qquad \forall k \in \mathbb{N}^*$$

On dit que Y suit une loi géométrique de paramètre p et on note $Y \sim \mathcal{G}(p)$. On montre que $\mathbb{E}(Y) = \frac{1}{p}$ et $\mathbb{V}(Y) = \frac{q}{p^2}$.

3.5 Loi géométrique généralisée ou loi binomiale négative

On considère une expérience aléatoire à deux issues : A (succès) et $\overline{A}$ (échec) avec $\mathbb{P}(A) = p$ et $\mathbb{P}(\overline{A}) = q = 1 - p$. On répète cette expérience de manière identique et indépendantes et on note T_r le nombre d'épreuve nécessaires pour obtenir le $r^{\text{ème}}$ succès. On a :

$$\mathbb{P}(T_r = k) = C_{k-1}^{r-1}(1-p)^{k-r}p^r \qquad \forall k \in \mathbb{N}^* \setminus \{1, 2, \ldots, r-1\}$$

On dit que T_r suit une loi binomiale négative de paramètre r et p et on note $T_r \sim \overline{\mathcal{B}}(r, p)$. On montre que $\mathbb{E}(Y) = \frac{r}{p}$ et $\mathbb{V}(Y) = \frac{rq}{p^2}$. On peut montrer que T_r est la somme de r variables géométriques indépendantes de paramètre p.

3.6 Loi hypergéométrique

Une urne contient N boules, dont N_1 noires et $N_2 = N - N_1$ blanches. On tire n boules sans les remettre et on note Z le nombre de boules noires obtenues. Z peut prendre toutes les valeurs entières entre $\max(o, n - N_2)$ et $\min(n, N_1)$ et pour k entre ces deux valeurs, on a :

$$\mathbb{P}(Z = k) = \frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}$$

On dit que Z suit une loi hypergéométrique de paramètres N, n, p où $p = \frac{N_1}{N}$ est la proportion de boules noires dans l'urne, et on note $Z \sim \mathcal{H}(N, n, p)$.

4 Exercices

4.1 *Calcul*

1. Calculer
   (a) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 & -1 \\ -1 & -2 & -3 \end{pmatrix}$
   (b) $i \begin{pmatrix} 1 & i & 2 & -i \\ 3 & 4 & -2 & i \\ 3 & 4 & -2 & i \\ i & -1 & 0 & 2 \end{pmatrix}$
   (c) $\begin{pmatrix} 3 & -5 & 4 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 6 & 5 & 0 \\ 8 & -2 & 4 \\ 0 & 3 & 1 \\ 1 & 0 & 2 \end{pmatrix}$
   (d) $\begin{pmatrix} 5 & 2 \\ 3 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$
2. Soient A, B et C les matrices définies par :

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 & 0 & 2 \\ -1 & -6 & 7 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

1. Déterminer une matrice X telle que : $A + 3B - 2X = 0$
2. Calculer AC et CA

4.2 *Puissance*

Soit M, P, D les matrices définies par ;

$$M = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix} \qquad P = \begin{pmatrix} -4 & -4 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Montrer la relation $M = PDP^{-1}$ et en déduire M^5

4.3 *Voisins*

1. Mon voisin a deux enfants dont une fille. Quelle est la probabilité que l'autre soit un garçon
2. Mon autre voisin a deux enfants. Le plus jeune est une fille. Quelle est la probabilité que l'aîné soit un garçon ?

4.4 *Parapluie*

On cherche un parapluie qui, avec la probabilité $p/7$, se trouve dans l'un quelconque des 7 étages d'un immeuble. On a exploré en vain les 6 premiers étages. Quelle est la probabilité que le parapluie se trouve au 7ème étage ? Soit $f(p)$ cette probabilité. Représenter graphiquement $f(p)$.

4.5 *Moutons*

On a décelé dans un élevage de moutons une probabilité 0.3 pour qu'un mouton soit atteint par une maladie M. La probabilité qu'un mouton qui n'est pas atteint par M ait une réaction négative à un test est 0.9. S'il est atteint par M, la probabilité qu'il ait une réaction positive à T est 0.8. Quelle est la probabilité qu'un mouton pris au hasard et ayant une réaction positive soit atteint par M ?

4.6 *Dur métier*

Une princesse est retenue prisonnière dans un château. Un prince charmant se présente pour essayer de la délivrer. Il entre dans la grande cour du château et, se trouvant alors devant trois portes, il en choisit une au hasard (avec équiprobabilité). Or, s'il ouvre la première porte, il découvre la princesse et parvient a la délivrer ; s'il ouvre la seconde, il tombe nez à nez avec un dragon qui le dévore ; et s'il ouvre la troisième, il rencontre une vieille sorcière qui le reconduit au dehors après l'avoir forcé à boire un philtre qui fait perdre au prince la mémoire de ce qu'il a vu au château. Le prince renouvelle ses tentatives jusqu'à ce qu'il périsse ou qu'il réussisse à délivrer la princesse.

1. (a) Quelle la probabilité pour que le prince réussisse à délivrer la princesse ?
   (b) Quel est le nombre moyen de tentatives que le prince peut et doit effectuer ? (*Nous désignons bien sûr par "nombre moyen de ..." l'espérance mathématique du nombre de ..*)
2. Si le prince succombe avant d'avoir réussi à délivrer la princesse, un autre se présente aussitôt pour tenter d'accomplir cette mission et ainsi de suite...
   (a) Quel est le nombre moyen de tentatives qui sont nécessaires pour parvenir à délivrer la princesse ?
   (b) Quel est le nombre moyen de princes qui sont dévorés par le dragon au cours de ces tentatives ?
   (c) Calculer la probabilité que le nombre de princes qui périssent ainsi soit strictement supérieur à dix.

Chapitre 2

Couple de variables aléatoires, loi marginale, loi conditionnelle

1 Couple de variables aléatoires

Soient X et Y deux variables aléatoires discrètes. La *loi du couple* (X, Y) est la donnée de :

$$\{\mathbb{P}((X,Y)=(x,y)) \text{ , } (x,y) \text{ valeurs possibles de } (X,Y)\}$$

Remarque : $\mathbb{P}((X,Y)=(x,y)) = \mathbb{P}(X=x, Y=y)$
Si $\phi\ : \mathbb{R}^2 \Rightarrow \mathbb{R}$ alors $\mathbb{E}(\phi(X,Y)) = \sum_x \sum_y \phi(x,y)\mathbb{P}(X=x,Y=y)$.
En particulier : $\mathbb{E}(XY) = \sum_x \sum_y xy\mathbb{P}(X=x,Y=y)$.

1.1 Espérance du couple (X, Y)

$$\mathbb{E}(X,Y) = (\mathbb{E}(X), \mathbb{E}(Y))$$

1.2 Matrice de variance-covariance du couple (X, Y)

$$\Gamma(X,Y) = \begin{pmatrix} \mathbb{V}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \mathbb{V}(Y) \end{pmatrix}$$

avec $\text{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
coefficient de corrélation : $\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$
Remarque : $\text{Cov}(X,Y)$ et $\rho_{X,Y}$ mesurent la dépendance linéaire entre X et Y :

- X et Y indépendantes $\Rightarrow \text{Cov}(X,Y) = 0$, $\rho_{X,Y} = 0$
- $\rho_{X,Y} = 1 \Rightarrow$ il existe α et β tel que $Y = \alpha X + \beta$

2 Loi marginale

Si on connaît la loi du couple (X, Y), on peut retrouver la loi de X et celle de Y. La loi de X est alors appelée *loi marginale de X* :

$$\mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y)$$

Par contre, si on connaît les deux lois marginales, on ne peut pas en général reconstituer la loi du couple, sauf si X et Y sont indépendantes :

Théorème 2.1 *X et Y indépendantes* $\Leftrightarrow \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$. *C'est-à-dire que si les variables sont indépendantes, la loi du couple est égale au produit des lois marginales.*

3 Loi conditionnelle

Si (X, Y) est un couple de variables aléatoires discrètes, on appelle *loi conditionnelle de X sachant $Y = y$* la donnée de :

$$\{\mathbb{P}((X = x|Y = y) \text{ , } x \text{ valeur possible de } X\}$$

Rappel : $\mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(Y=y)}$
Si on connaît la loi conditionnelle de X sachant $Y = y$ pour toutes les valeurs possibles de Y, on dit qu'on connaît la loi conditionnelle de X sachant Y.
Remarque : À partir de la loi de Y et de la loi conditionnelle de X sachant Y, on retrouve la loi du couple (X, Y) :

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x|Y = y)\mathbb{P}(Y = y)$$

Théorème 2.2 *X et Y indépendantes* $\Leftrightarrow \mathbb{P}(X = x|Y = y) = \mathbb{P}(X = x)$. *C'est-à-dire que si les variables sont indépendantes, la loi conditionnelle de X sachant Y est égale à la loi marginale de X.*

4 Espérance conditionnelle

$\mathbb{P}(X = x|Y = y)$ est une probabilité. On peut donc définir l'espérance de X sachant $Y = y$:

$$\mathbb{E}(X = x|Y = y) = \sum_x x\mathbb{P}(X = x|Y = y)$$

Si on connaît la loi conditionnelle de X sachant Y, on connaît $\mathbb{E}(X = x|Y = y)$ pour toutes les valeurs possibles y de Y.
On peut alors définir la variable $\mathbb{E}(X|Y)$, qui vaut $\mathbb{E}(X|Y = y)$ quand $Y = y$.
Cette variable s'appelle l'*espérance conditionnelle de X sachant Y*, elle prend la valeur $\mathbb{E}(X|Y = y)$ avec la probabilité $\mathbb{P}(Y = y)$.
On peut alors montrer les propriétés suivantes :

1. $\mathbb{E}(\mathbb{E}(X|Y=y)) = \mathbb{E}(X)$
2. Si X et Y sont indépendantes , alors $\mathbb{E}(X|Y) = \mathbb{E}(X)$
3. $\mathbb{E}(\phi(X)\psi(Y)|Y) = \psi(Y)\mathbb{E}(\phi(X)|Y)$.
   En particulier $\mathbb{E}(XY|Y) = Y\mathbb{E}(X|Y)$

5 Exercices

5.1 Lois marginales et conditionnelles

X et Y étant deux variables aléatoires discrètes ayant pour ensemble de valeurs possibles $\mathcal{X} = \{1, 2, 3\}$ et $\mathcal{Y} = \{-1, 1\}$, la distribution de probabilité du couple (X, Y) est donnée par le tableau suivant :

$y \backslash x$	1	2	3
-1	0.1	0.3	0.1
1	0.2	0.1	0.2

1. Quelles sont les lois de probabilité marginales des variables aléatoires X et Y ? Calculer $\mathbb{E}(X)$ et $\mathbb{E}(Y)$.
2. Quelle est la distribution de probabilité conditionnelle de la variable aléatoire X ? Calculer $\mathbb{E}(X|Y)$. (*i.e.* $\mathbb{E}(X|Y = 1)$ et $\mathbb{E}(X|Y = -1)$).
3. En déduire $\mathbb{E}(XY)$. Indication : on utilisera la propriété de l'espérance conditionnelle : $\mathbb{E}(XY|Y) = Y\mathbb{E}(X|Y)$.
4. X et Y sont-elles indépendantes ?

5.2 Binomiale et Poisson

1. Soient X et Y deux variables aléatoires indépendantes de loi de Poisson $\mathcal{P}(\lambda_1)$ et $\mathcal{P}(\lambda_2)$.
   (a) On pose $S = X + Y$. Quelles est la loi de S ?
   (b) Montrer que la loi conditionnelle de X sachant S est binomiale :
   $$\mathbb{P}(X = k|S = s) = C_s^k p^k (1 - p)^{s-k} \quad \text{avec } p = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$
2. Soient X et Y deux variables aléatoires indépendantes de loi binomiale $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$.
   (a) On pose $S = X + Y$. Quelle est la loi de S ?
   (b) Montrer que la loi conditionnelle de X sachant S est hypergéométrique.

5.3 Famille nombreuse

On admet que le nombre d'enfants dans une famille est une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$. On considère qu'à chaque naissance la probabilité d'avoir un garçon est $1/2$.

1. Exprimer la probabilité d'observer k garçons dans une famille dont on sait qu'elle contient n enfants.
2. Déterminer la loi de probabilité du nombre de garçons que l'on peut observer dans une famille.

Chapitre 3

Chaînes de Markov

1 Définitions

Soient $X_0, X_1, X_2, \ldots, X_n, \ldots$ une suite de variables aléatoires définies sur un même espace de probabilité Ω et à valeurs dans un même espace E :

$$X_n \ : \Omega \mapsto E$$

Dans tout ce cours, E sera fini ou dénombrable. On peut donc l'identifier à $\mathbb{N}$ ou à une partie de $\mathbb{N}$. On dit que E est l'espace d'états.
Exemple :

$$\begin{array}{rcl} X_n & : & \text{temps qu'il fait au jour } n \\ E & = & \{\text{pluie, beau temps, neige}\} \\ & \simeq & \{1, 2, 3\} \end{array}$$

$$\begin{array}{ll} X_n = 1 & \text{si le } n^{\text{ème}} \text{ jour est pluvieux} \\ X_n = 2 & \text{si le } n^{\text{ème}} \text{ jour est beau} \\ X_n = 3 & \text{si le } n^{\text{ème}} \text{ jour est neigeux} \end{array}$$

Définition 3.1 *On dit que la suite $(X_n)_{n\in\mathbb{N}}$ est une* chaîne de Markov *si :*

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \ldots, X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) \quad \forall i_0, i_1, \ldots, i_n, i_{n+1} \in E$$

Le futur ne dépend que du présent et pas du passé. On peut aussi dire que les variables sont indépendantes mais de mémoire 1 seulement.

Définition 3.2 *Si $(X_n)_{n\in\mathbb{N}}$ est une* chaîne de Markov*, on dit qu'elle est* homogène *si :*

$$\forall i, j \in E \quad \forall n \in \mathbb{N} \quad \mathbb{P}(X_{n+1} = j | X_n = i) = \pi(i, j)$$

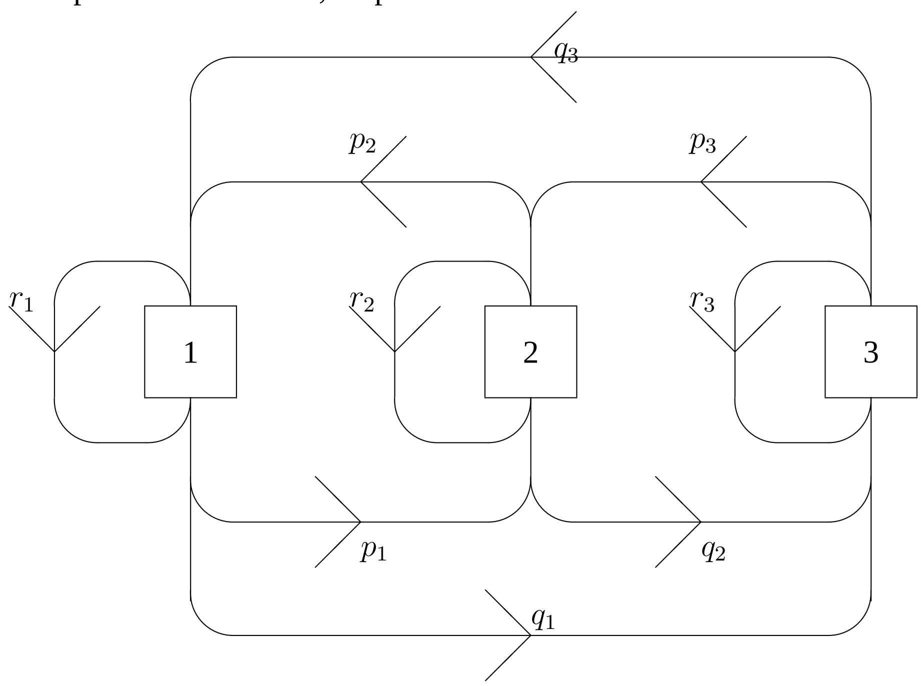
La probabilité de passer de l'état i à l'état j au $n^{\text{ème}}$ coup ne dépend que de i et de j et pas de n.

Exemple : $\mathbb{P}(X_4 = 2|X_3 = 3) = \mathbb{P}(X_{11} = 2|X_{10} = 3) = \mathbb{P}(X_1 = 2|X_0 = 3) = \pi(3, 2)$
Nous ne nous intéresserons qu'aux chaînes de Markov homogènes d'espace d'état fini ou dénombrable.
Les probabilités $\pi(i, j)$ s'appellent *probabilités de transition* ou *probabilités de passage*.
$\pi(i, j)$ est la probabilité de passer de l'état i à l'état j.
On appelle *matrice de transition* et on note Π la matrice $\Pi = (\pi)_{i,j\in E}$. Si l'espace d'état E est infini, c'est un abus de langage.
Exemple : On pose

1	:	pluie
2	:	beau temps
3	:	neige

$$\Pi = \begin{pmatrix} r_1 & p_1 & q_1 \\ p_2 & r_2 & q_2 \\ q_3 & p_3 & r_3 \end{pmatrix}$$

Comme l'espace d'états est fini, on peut visualiser la chaîne de Markov sur un graphe.



La matrice de transition est une *matrice stochastique*,c'est-à-dire :
- $\forall i, j\ \pi(i, j) \geq 0$
- $\forall i \sum_j \pi(i, j) = 1$

Preuve :
- $\pi(i, j) = \mathbb{P}(X_{n+1} = j|X_n = i) \geq 0$
- $\sum_j \pi(i, j) = \sum_j \mathbb{P}(X_{n+1} = j|X_n = i) = 1$ car $\mathbb{P}(X_{n+1} = j|X_n = i)$ est une probabilité

2 Propriété fondamentale des chaînes de Markov homogènes

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mathbb{P}(X_0 = i_0)\pi(i_0, i_1)\pi(i_1, i_2)\dots\pi(i_{n-1}, i_n) \quad \forall i_0, i_1, \dots, i_n \in E$$

Preuve : On utilise la relation $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$

$$\begin{aligned}
&\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0)\\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \dots, X_0 = i_0)\mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0)\\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})\mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0)\\
&= \pi(i_{n-1}, i_n)\mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0)\\
&= \pi(i_{n-2}, i_{n-1})\pi(i_{n-1}, i_n)\mathbb{P}(X_{n-2} = i_{n-2}, \dots, X_0 = i_0)\\
&= \vdots\\
&= \mathbb{P}(X_0 = i_0)\pi(i_0, i_1)\pi(i_1, i_2)\dots\pi(i_{n-1}, i_n)
\end{aligned}$$

La loi de la chaîne ne dépend que de la matrice de transition Π et de la loi de X_0 (qu'on appelle la loi initiale).

3 Calcul de la loi de X_n

Pour simplifier, on suppose que l'espace d'états est fini :

$$E = \{1, 2, \dots, d\}$$

Il y a d états possibles
La loi de X_n est représentée par le vecteur ligne :

$$\mathbb{P}_n = (\mathbb{P}(X_n = 1), \mathbb{P}(X_n = 2), \dots, \mathbb{P}(X_n = d))$$

On utilise la formule de Bayes avec la partition disjointe :

$$(X_{n-1} = 1), (X_{n-1} = 2), \cdots, (X_{n-1} = d)$$

et donc :

$$\begin{aligned}
\mathbb{P}(X_n = 1) &= \sum_{i=1}^{d} \mathbb{P}(X_n = 1 | X_{n-1} = i)\mathbb{P}(X_{n-1} = i)\\
&= \sum_{i=1}^{d} \pi(i, 1)\mathbb{P}(X_{n-1} = i)
\end{aligned}$$

De même pour tout j :

$$\mathbb{P}(X_n = j) = \sum_{i=1}^{d} \pi(i, j)\mathbb{P}(X_{n-1} = i)$$

D'où la formule fondamentale :

$$\mathbb{P}_n = \mathbb{P}_{n-1}\Pi$$

on répète l'opération :

$$\begin{aligned}
\mathbb{P}_n &= \mathbb{P}_{n-1}\Pi \\
&= (\mathbb{P}_{n_2}\Pi)\Pi = \mathbb{P}_{n-2}\Pi^2 \\
&= (\mathbb{P}_{n_3}\Pi)\Pi = \mathbb{P}_{n-3}\Pi^3 \\
&= \vdots \\
&= \mathbb{P}_0\Pi^n
\end{aligned}$$

où $P_0 = (\mathbb{P}(X_0 = 1), \mathbb{P}(X_0 = 2), \ldots, \mathbb{P}(X_0 = d))$ est la *loi initiale* de la chaîne de Markov.

Théorème 3.1 $\mathbb{P}_n = \mathbb{P}_0\Pi^n$ *où* P_0 *est la loi initiale et où* Π *est la matrice de transition.*

Ceci implique qu'une chaîne de Markov homogène à espace d'états dénombrable est entièrement caractérisée par P_0 et Π.

4 Matrice de transition dans le temps

$\pi(i,j)$ est la probabilité de passer de i à j en 1 coup.

Théorème 3.2 *La probabilité de passer de* i *à* j *en* m *coups est égale à* $(\Pi^m)(i,j)$*, le coefficient de la* $i^{\text{ème}}$ *ligne ,* $j^{\text{ème}}$ *colonne de* Π^m*. On dit que* Π^m *est la matrice de transition en* m *coups.*

Remarques :

- $(\pi^m)(i,j) = \mathbb{P}(X_m = j|X_0 = i) = \mathbb{P}(X_{5+m} = j|X_5 = i) = \mathbb{P}(X_{k+m}|X_k = i)$
- $(\Pi^m)(i,j) \neq (\pi(i,j)^m)$

Preuve du théorème : On doit montrer que $\mathbb{P}(X_m = j|X_0 = i) = (\pi^m)(i,j)$. Par récurrence :

$m = 1$ **:** $\mathbb{P}(X_1 = j|X_0 = i) = \pi(i,j)$

m **:** Supposons $\mathbb{P}(X_m = j|X_0 = i) = (\pi^m)\,(i,j)$

$m+1$ **:**

$$\begin{aligned}
\mathbb{P}(X_{m+1} = j|X_0 = i) &= \frac{\mathbb{P}(X_{m+1} = j, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k=1}^{d} \frac{\mathbb{P}(X_{m+1} = j, X_m = k, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k=1}^{d} \frac{\mathbb{P}(X_{m+1} = j|X_m = k, X_0 = i)\mathbb{P}(X_m = k, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k=1}^{d} \frac{\mathbb{P}(X_{m+1} = j|X_m = k, X_0 = i)\mathbb{P}(X_m = k|X_0 = i)\ \cancel{\mathbb{P}}(X_0 = i)}{\cancel{\mathbb{P}}(X_0 = i)}
\end{aligned}$$

$$\begin{aligned} &= \sum_{k=1}^{d} \pi(k,j)\left(\pi^m\right)(i,k) \\ &= \left(\pi^{m+1}\right)(i,j) \end{aligned}$$

Remarques : (π^m) est encore une matrice stochastique car :

- $\forall i, j$, $(\pi^m)(i,j) \geq 0$
- $\sum_{i=1}^{d} (\pi^m)(i,j) = 1 \; \forall i$

5 Exercices

5.1 Weather in the Land of Oz

In the Land of Oz they never have two nice days in a row. If they have a nice day they are just as likely to have snow as rain in the next . If they have snow (or rain) they have an even chance of having the same in the next day. If there is a change from snow or rain, only half of the time is this change to a nice day.

1. Montrer que le temps au pays d'Oz est une chaîne de Markov homogène et donner sa matrice de transition Π. Calculer Π^5.

   Vous avez prévu d'arriver au pays d'Oz un lundi mais pas dans n'importe quelles conditions : vous n'entreprenez le voyage qui s'il a plu au pays d'Oz le dimanche sinon vous repoussez votre voyage à la semaine suivante.
2. Quelle est la probabilité qu'il ne fasse pas beau le lundi de votre arrivée ?
3. Quelle est la probabilité qu'il fasse beau le samedi qui suit votre arrivée ?
4. Vous arrivez enfin et il neige. Quelle est la probabilité qu'il fasse beau le samedi ?
5. Le mardi vous dormez toute la journée et le mercredi il neige encore. Quelle est la probabilité qu'il ait fait beau pendant votre sommeil ?

5.2 Propriétés des chaînes de Markov homogènes

Montrer que si $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène alors :

1. Pour tous états $x_0, x_1, \ldots, x_n$ on a :

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_0 = x_0)\pi(x_0, x_1)\pi(x_1, x_2)\cdots\pi(x_{n-1}, x_n)$$

2. Pour tous états $x, y \in E$ on a :

$$P(X_{n+k} = y|X_n = x) = (\pi^k)(x, y)$$

   où $(\pi^k)(x, y)$ est la $x^{\text{ème}}$ ligne, $y^{\text{ème}}$ colonne de la matrice Π^k.

5.3 Chaîne à deux états

Soit une chaîne de Markov à deux états 0 et 1 et de loi initiale P_0 et de matrice de transition :

$$\Pi = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \qquad 0 \leq p, q \leq 1$$

1. Calculer $\mathbb{P}(X_1 = 0|X_0 = 0 \text{ et } X_2 = 0)$ et $\mathbb{P}(X_1 \neq X_2)$
2. Calculer $\mathbb{P}_n$ en fonction de $\mathbb{P}_{n-1}$. Puis $\mathbb{P}_n$ en fonction de $\mathbb{P}_0$. En déduire Π^n.

5.4 *Météo*

1. On suppose que le fait qu'il pleuve demain ou non dépend seulement du fait qu'il pleuve ou non aujourd'hui. On suppose que si il pleut aujourd'hui, il pleuvra demain avec une probabilité $1-\alpha$, et s'il ne pleut pas aujourd'hui, il pleuvra demain avec une probabilité β. On représente le temps qu'il fait par 0 quand il pleut et 1 quand il ne pleut pas. Ceci définit un processus X_n état du temps au jour n.
   (a) Écrire la matrice de transition Π.
   (b) Calculer la probabilité qu'il pleuve dans quatre jours sachant qu'il pleut aujourd'hui.
   (c) Calculer la probabilité qu'il pleuve quatre jours de suite sachant qu'il pleut aujourd'hui.
   (d) Montrer qu'il existe une probabilité invariante Q, la calculer en fonction de α et β
   (e) Convergence ?
2. On suppose maintenant que le fait qu'il pleuve ou non demain dépend aussi du temps qu'il a fait hier.
   - S'il a plu hier et aujourd'hui, il pleuvra demain avec une probabilité $1-\alpha$.
   - S'il a plu aujourd'hui et pas hier, il pleuvra demain avec une probabilité $1-\gamma$.
   - S'il a plu hier et pas aujourd'hui, il pleuvra demain avec une probabilité β.
   - S'il n'a plu ni hier ni aujourd'hui, il pleuvra demain avec une probabilité δ.

   (a) Si l'on considère le processus défini à la question 1, montrer que ce processus n'est plus une chaîne de Markov (sauf conditions particulières sur $\alpha, \beta, \gamma, \delta$).
   (b) On définit alors le processus Y_n ainsi :
   - $Y_n = 0$ s'il pleut au jour n et au jour $n-1$.
   - $Y_n = 1$ s'il pleut au jour n et pas au jour $n-1$.
   - $Y_n = 2$ s'il pleut au jour $n-1$ et pas au jour n.
   - $Y_n = 3$ s'il ne pleut ni au jour n ni au jour $n-1$.

   Écrire la matrice de transition
   (c) Sachant qu'il a plu lundi et mardi, calculer la probabilité qu'il pleuve jeudi.
   (d) Sachant qu'il n'a plu ni samedi ni dimanche, calculer la probabilité qu'il pleuve mardi.
   (e) Déterminer la probabilité stationnaire et la proportion de jours où il pleut, observée sur un temps assez long, en supposant que $\alpha, \beta, \gamma, \delta \in]0, 1[$.

Chapitre 4

Convergence des chaînes de Markov homogènes à espace d'états dénombrable

Question : peut-on évaluer $\mathbb{P}(X_n = j)$? (sans connaître n)
On a $\mathbb{P}_n = (\mathbb{P}(X_n = 1), \mathbb{P}(X_n = 2), \ldots, \mathbb{P}(X_n = d))$. A priori $P_n = P_0\Pi^n$ dépend de n.
Exemple : (pays d'Oz)

$$\begin{array}{rcl} \mathbb{P}_0 & = & (1,0,0) \\ \mathbb{P}_1 & = & \mathbb{P}_0\Pi = (1/2,\, 1/4,\, 1/4) \\ \mathbb{P}_2 & = & \mathbb{P}_1\Pi = (7/16,\, 3/16,\, 6/16) \end{array}$$

À première vue, la question : "Quelle est la probabilité qu'il fasse beau au pays d'Oz ?", c'est-à-dire "Combien vaut $\mathbb{P}(X_n = 2)$?" n'a pas de réponse. Mais si n devient grand :

$$\begin{array}{rcl} \mathbb{P}_3 & = & (0.4063, 0.2031, 0.3906) \\ \mathbb{P}_4 & = & (0.4023, 0.1992, 0.3984) \\ \mathbb{P}_5 & = & (0.4004, 0.2002, 0.3994) \\ \mathbb{P}_6 & = & (0.4, 0.2, 0.4) \\ \mathbb{P}_7 & = & (0.4, 0.2, 0.4) \end{array}$$

On peut répondre à la question si $\mathbb{P}(X_n = j) \overset{n\to+\infty}{\longrightarrow} Q(j)$
La probabilité qu'il fasse beau au pays d'Oz est 0.2

Définition 4.1 *On dit que la chaîne $(X_n)_{n\in\mathbb{N}}$ converge si*

$$\mathbb{P}(X_n = j) \overset{n\to+\infty}{\longrightarrow} Q(j)$$

pour tous les états j.

Si on note $Q = (Q(1), Q(2), \ldots, Q(d))$ alors :

$$(X_n)_{n\in\mathbb{N}} \text{ converge } \Leftrightarrow \mathbb{P}_n \overset{n\to+\infty}{\longrightarrow} Q$$

1 Probabilité invariante et convergence

On va chercher une condition nécessaire de convergence.
Si $\mathbb{P}_n \overset{n\to+\infty}{\longrightarrow} Q$ alors $\mathbb{P}_n = \mathbb{P}_{n-1}\Pi$ implique $Q = Q\Pi$.
On dit que Q est une probabilité invariante ou stationnaire.
condition nécessaire de convergence : existence d'une probabilité invariante Q.

Théorème 4.1

$$\forall \mathbb{P}_0 \qquad \mathbb{P}_n \overset{n\to+\infty}{\longrightarrow} Q \Leftrightarrow \Pi^n \overset{n\to+\infty}{\longrightarrow} \Gamma = \begin{pmatrix} Q(1) & Q(2) & \cdots & Q(d) \\ Q(1) & Q(2) & \cdots & Q(d) \\ \vdots & \vdots & & \vdots \\ Q(1) & Q(2) & \cdots & Q(d) \end{pmatrix}$$

et dans ce cas $Q = Q\Pi$

Preuve :

- $\mathbb{P}_n = \mathbb{P}_0\Pi^n$ donc si $\mathbb{P}_n$ converge alors $\Pi^n \overset{n\to+\infty}{\longrightarrow} \Gamma$.
  $(0,\ldots,0,i,0,\ldots,0)\Pi^n \overset{n\to+\infty}{\longrightarrow} (0,\ldots,0,i,0,\ldots,0)\Gamma = (\Gamma_{1,i},\Gamma_{2,i},\ldots,\Gamma_{d,i}) = Q$
- Réciproquement : si Π^n converge vers un Γ de cette forme :

$$\mathbb{P}_n = \mathbb{P}_0\Pi^n \to \mathbb{P}_0\Gamma$$

  et

$$\mathbb{P}_0\Gamma = \left(\sum_{i=1}^{d}\mathbb{P}_0(i)Q(1), \sum_{i=1}^{d}\mathbb{P}_0(i)Q(2),\ldots,\right) = (Q(1),Q(2),\ldots)$$

  car $\sum_{i=1}^{d}\mathbb{P}_0(i) = 1$.

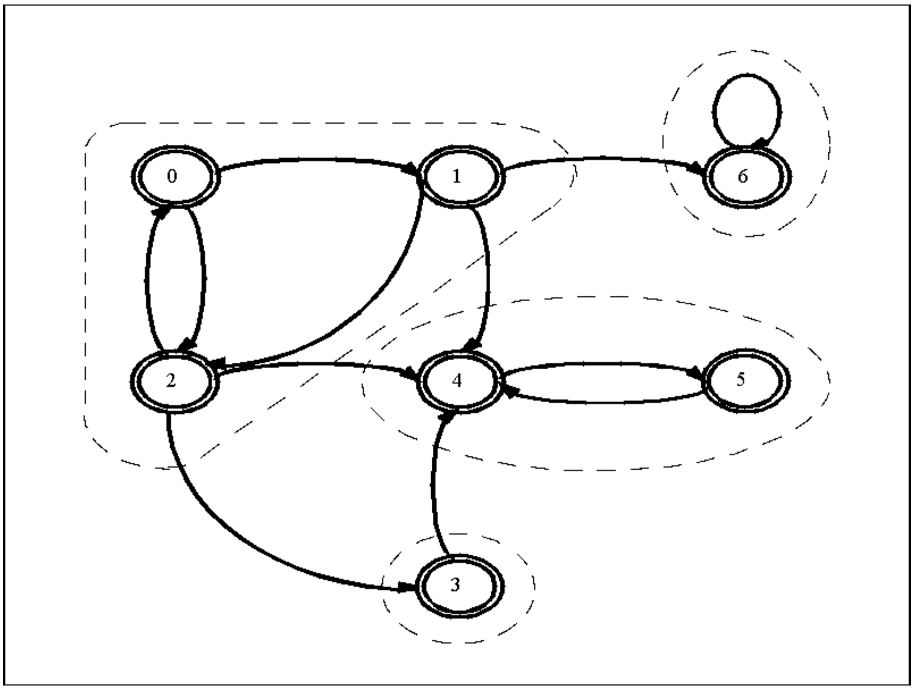
Théorème 4.2 *Si l'espace d'état est fini et si* Π *a une seule valeur propre de module 1 (elle en a toujours au moins une), alors* $\Pi^n \overset{n\to+\infty}{\longrightarrow} \begin{pmatrix} \ldots & Q & \ldots \\ \ldots & Q & \ldots \\ & \ldots & \\ \ldots & Q & \ldots \end{pmatrix}$ *et donc* $\mathbb{P}_n \overset{n\to+\infty}{\longrightarrow} Q$

Q est le vecteur propre associé à la valeur propre de module 1

2 Classification des chaînes de Markov

Dans la définition que nous avons donnée d'une chaîne de Markov, l'évolution du processus au cours du temps à partir d'un état donné est entièrement décrite par la matrice des probabilités de transition. On peut aussi voir une chaîne de Markov comme un ensemble d'états entre lesquels s'effectuent des transitions. Certaines transitions sont possibles (probabilité de transition strictement positive) alors que d'autres sont impossibles (probabilité de transition nulle). Ceci nous amène à vouloir visualiser une chaîne de Markov en représentant chaque état par un sommet et chaque transition par un arc. Il faut

Figure 4.1 – graphe des transitions possibles d'une chaîne de Markov



noter qu'un arc possède une orientation. Ce point de vue structurel consiste en fait à visualiser le graphe des transitions possibles d'une chaîne de Markov (voir figure 4.1). Dans la mesure où les arcs sont orientés, on parle de **graphe orienté**. Si des poids sont associés aux arcs, on parle **d'automate**

Dans la théorie des graphes, on appelle **chemin** une succession d'arcs, telle que l'extrémité du $n^{\text{ème}}$ arc soit l'origine du $(n+1)^{\text{ème}}$ arc et on appelle **circuit** un chemin fermé. Le graphe des transitions possibles de la figure 4.1 comporte par exemple le chemin [0,1,2,3,4,5] et le circuit [0,1,2,0].

Pour la suite on note $\pi_n(i,j)$ la probabilité que le système soit dans l'état i au temps t et dans l'état j au temps $t+n$:

$$\pi_n(i,j) = \mathbb{P}(X_{t+n} = j | X_t = i) = \mathbb{P}(X_n = j | X_0 = i)$$

On dit que l'état j est **accessible** à partir de l'état i si la probabilité de passer de i à j est non nulle :

$$i \to j \Longleftrightarrow \exists n \geq 0 : \pi_n(i,j) > 0$$

En théorie des graphes ceci signifie qu'il existe un chemin entre i et j.

On dit que les états i et j **communiquent** si chacun d'eux est accessible à partir de l'autre :

$$i \leftrightarrow j \Longleftrightarrow \begin{cases} i \to j \\ j \to i \end{cases}$$

Pour que deux états ne communiquent pas il faut que l'un des deux ne soit pas accessible à partir de l'autre, c'est-à-dire :

$$\forall n \geq 0 \quad \pi_n(i,j) = 0 \text{ ou } \quad \forall n \geq 0 \pi_n(j,i) = 0$$

La relation de communication entre deux états est réflexive (par convention $\forall i \; \pi_0(i,i) = 1$), symétrique (par définition) et transitive, c'est donc une **relation d'équivalence**.
Il est donc possible de construire une partition des états d'une chaîne de Markov en classes d'équivalence telle que tous les états d'une classe communiquent entre eux et que deux états appartenant à deux classes différentes ne communiquent pas. Par construction, ces classes sont deux à deux disjointes et leur réunion est l'ensemble des états.
En théorie des graphes, une classe d'équivalence correspond à une **composante fortement connexe**, c'est-à-dire dont tous les éléments sont communiquants. On peut donc construire le **graphe réduit** (par exemple la figure 4.2). Dans ce graphe, les sommets représentent les classes et les arcs représentent les transitions possibles entre classes. Ce graphe possède la propriété d'être sans **circuit** (on ne peut jamais revenir au point d'origine), tous les circuits du graphe d'origine des transitions possibles ayant servi à construire les différentes classes.
Il est alors possible de distinguer deux types de classe :

- une classe est dite **transitoire** s'il est possible d'en sortir mais dans ce cas, le processus ne pourra plus jamais y revenir (classe (0,1,2) et classe (3) dans la figure 4.2) ;
- une classe est dite **récurrente** s'il est impossible de la quitter (classe (4,5) et classe (6) dans la figure 4.2).

Si une classe récurrente est composée d'un seul état, cet état est dit absorbant (état 6 dans la figure 4.2). Un état i absorbant est donc tel qu'une fois dans cet état on ne peut le quitter (par exemple la ruine dans le cas du jeu de "Pile ou Face"). En terme de probabilités de transition, ceci signifie que $\forall k \neq i$, $\pi_{ik} = 0$ et donc $\pi_{ii} = 1$.
Les états absorbants sont très particuliers puisqu'ils constituent des états terminaux du système. Il est notamment intéressant d'étudier les probabilités d'absorption, c'est-à-dire les probabilités que le système finisse par atteindre un tel état.
Les états d'une classe transitoire sont dits transitoires alors que les états d'une classe récurrente sont dits récurrents. Un état absorbant est donc un type particulier d'état récurrent.
Une chaîne de Markov pour laquelle il n'existe qu'une seule classe récurrente (égale à l'ensemble des états) est dite **irréductible**. Ceci signifie que tous les états communiquent.

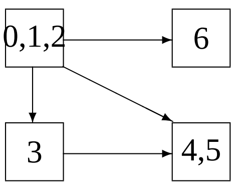


Figure 4.2 – exemple de graphe réduit

Pour un état i de la chaîne, on appelle **temps de retour** le temps minimal pour revenir à l'état i ; c'est-à-dire le plus petit n tel que $\pi_n(i,i) > 0$.

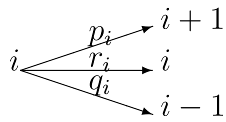


Figure 4.3 – Exemple de marche aléatoire

Soit i un état d'une chaîne de Markov. La **période de retour** de i, notée T_i est la quantité définie par :

$$\begin{array}{lcl} \text{Si } n = kT_i, k \in \mathbb{N} & \Rightarrow & \pi_n(i,i) > 0 \\ \text{Si } n \neq kT_i, k \in \mathbb{N} & \Rightarrow & \pi_n(i,i) = 0 \end{array}$$

c'est-à-dire que les retours à l'état i ne sont possibles que pour des durées multiples à la période.
Une autre manière équivalente de dire les choses est de définir la période comme le pgcd$\{n \in \mathcal{T} : \pi_n(i,i) > 0\}$
L'état i est dit **périodique** si $T_i > 1$ et **apériodique** si $T_i = 1$.
Il est possible de montrer que deux états communiquants ont la même période et donc que la période est constante à l'intérieur des classes de communication.
La période commune des éléments de la classe est appelée **période de la classe**.
Si la chaîne est irréductible et qu'elle a une période, on parle d'une **chaîne périodique** ; si elle n'a pas de période on parle de **chaîne apériodique**.
Une chaîne irréductible et apériodique est dite **ergodique**.

Exemple 4.1 *Soit la matrice de transition de la chaîne :*

$$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

la période de la chaîne est 2, le système agit comme un métronome.

2.1 Exemple : marche aléatoire

Un individu se déplace dans une direction fixe et peut, à chaque étape, soit faire un pas en avant (avec une probabilité p_i), soit faire un pas en arrière (probabilité q_i), soit rester sur place (probabilité $r_i = 1 - p_i - q_i$).
On suppose que ce processus est homogène, ce qui signifie que les probabilités des trois événements dépendent de l'endroit i où l'individu se trouve mais pas de l'étape n. En

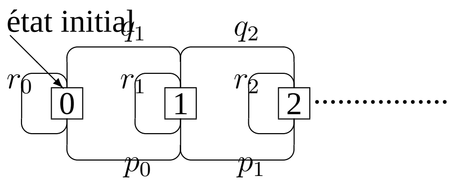


Figure 4.4 – Graphe d'une marche aléatoire

notant 0 le premier état, on obtient donc la matrice de transition :

$$\Pi = \begin{pmatrix} r_0 & p_0 & 0 & \dots & & \\ q_1 & r_1 & p_1 & 0 & \dots & \\ 0 & q_2 & r_2 & p_2 & 0 & \dots \\ \vdots & 0 & q_3 & r_3 & \ddots & \ddots \\ & \vdots & 0 & \ddots & \ddots & \\ & & \vdots & \ddots & & \end{pmatrix}$$

Cette matrice est de dimension finie ou infinie.
On peut représenter cette matrice sous forme d'un automate (voir figure 4.4)
De nombreux cas de marche aléatoire sont utilisés : fortune du joueur au jeu de "Pile ou Face", etc.

3 Résultat fondamental pour les chaînes à espace d'états fini

Théorème 4.3 • *Si la chaîne est irréductible alors il existe une unique probabilité invariante Q.*
- *Si la chaîne est irréductible et apériodique alors elle converge $\mathbb{P}_n \overset{n\to+\infty}{\longrightarrow} Q$ et la loi limite ne dépend pas de la loi initiale.*
- *Si la chaîne est irréductible et périodique de période d, alors :*
  - *pour chaque couple d'états i et j, il existe $0 \leq r < d$ tel que*

$$(\Pi^n)(i,j) = \begin{cases} = 0 & \text{si } n \neq md + r \\ > 0 & \text{si } n = md + r \end{cases}$$

  - *on a des limites différentes :*

$$\lim_{\substack{n=md+r \\ n\to+\infty}} \Pi^n = A_r$$

  pour tous les r entre 0 et $d-1$.
  Les coefficients $A_r(i,j)$ sont tous égaux à 0 ou $d \times Q(j)$. Dans ce cas la loi limite dépend de la loi initiale.
  En résumé si on se trouve sur la période il y a convergence et sinon il y a divergence.

4 Exercices

4.1 *Weather in the Land of Oz*

Sur un temps assez long, peut-on prévoir la proportion de jours de neige, de pluie et de soleil au pays d'Oz ? Dans l'affirmative faut-il attendre très longtemps pour que ces prévisions soient bonnes ?

4.2 *Chaîne produit*

Soient $X_1, \dots, X_n, \dots$ des variables aléatoires indépendantes et de même loi :

$$\mathbb{P}(X_i = -1) = p \quad \mathbb{P}(X_i = +1) = 1 - p$$

avec $0 \le p \le 1$
On pose $Z_0 = 1$ et pour tout entier $n \ge 1$ on pose $Z_n = \prod_{i=1}^{n} X_i$.

1. Quel est l'espace d'états de la chaîne Z_n ? Montrer que $(Z_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène. Déterminer sa matrice de transition et sa loi initiale. Les variables aléatoires $Z_0, Z_1, Z_2, \dots$ sont-elles indépendantes entre elles ? (la réponse dépend de la valeur de p.)
2. Établir que les probabilité de transition dans le temps (*i.e.* $\mathbb{P}(Z_{n+k} = j | Z_k = i)$) sont données par la matrice :

$$\begin{pmatrix} 1 - p_n & p_n \\ p_n & 1 - p_n \end{pmatrix}$$

avec $p_n = \frac{1-(1-2p)^n}{2}$
3. Déduire du 2 que la variable aléatoire Z_n possède une loi limite quand $n \to +\infty$ et déterminer cette loi limite. Établir ensuite que la suite de variables aléatoires (Z_n, Z_{n+1}) converge en loi et déterminer sa loi limite dans le cas où $p \in]0, 1[$.

4.3 *Reproduction diploide*

On considère une population diploide à deux allèles A et a et à générations non recouvrantes. L'objectif est de déterminer les fréquences limites des divers génotypes AA, Aa, aa dans le cas d'accouplement entre frère et sœur. Pour cela, on va étudier les fluctuations au cours du temps des fréquences des types d'accouchements.
Pour cette populations il existe six types distincts de couple : $E_1 = AA \times AA$, $E_2 = AA \times Aa$, $E_3 = Aa \times Aa$, $E_4 = Aa \times aa$, $E_5 = AA \times aa$, $E_6 = aa \times aa$. Les couples vont donner des descendants du type $b_1 = AA$, $b_2 = Aa$ et $b_3 = aa$.

1. Donner les probabilité $\mathbb{P}(b_1|E_i)$, $\mathbb{P}(b_2|E_i)$, $\mathbb{P}(b_3|E_i)$; $i = 1, \dots, 6$. On présentera les résultats sous forme de tableau.
2. Parmi les descendants directs d'un couple C_0 on fait des croisements frère-sœur : on choisit au hasard des individus de sexe opposé pour recréer C_1.

(a) Expliquer pourquoi on définit ainsi une chaîne de Markov

(b) Calculer la matrice de transition $\Pi = (\pi_{i,j})_{1\leq i,j\leq 6}$ où $\pi_{i,j}$ est la probabilité qu'un couple frère-sœur soit du type E_j alors que les parents sont du type E_i.

(c) Faire le graphe des états. Commenter. Quelles évolution prévoyez-vous pour la chaîne ?

3. Soit L_n le vecteur de la loi des croisements à la n-ième génération :

$$L_n = (\mathbb{P}(C_n = E_1), \mathbb{P}(C_n = E_2), \ldots, \mathbb{P}(C_n = E_6))$$

(a) Quelle relation de récurrence vérifie L_n ? Donner l'expression de L_n en fonction de L_0 et de Π.

(b) Déduire du graphe des états que Π^n converge quand n tend vers $+\infty$ vers une matrice Γ que l'on déterminera. En déduire $\lim_{n\to+\infty} L_n$ en fonction de L_0.

(c) Trouver la limite de L_n quand $n \to +\infty$ dans les trois cas suivants :
- L_0 est une loi déterministe en un croisement E_i $(i = 1, \ldots, 6)$
- Le couple initial a été choisi au hasard dans une population où la proportion de gène A est ½
- Le couple initial a été choisi au hasard parmi les individus rouges de cette population (AA : rouge, Aa : rouge, aa : blanc).

(d) Quelles sont les fréquences limites des génotypes AA, Aa, aa ?

4.4 *Chaîne de Wright*

Chaque population $\mathcal{F}_n$ (n est l'indice de génération) est formée de $2N$ gènes A et a. On considère une reproduction (haploide) de type aléatoire pur : on tire dans $\mathcal{F}_n$ avec répétition, $2N$ gènes pour former $\mathcal{F}_{n+1}$. Les $2N$ tirages sont indépendants. On note Y_n le nombre de gènes A dans la population $\mathcal{F}_n$.

1. Quelle est la loi de Y_{n+1} sachant $Y_n = j$? Calculer son espérance.
2. Montrer que Y_n est une chaîne de Markov. Quelle est la nature des états de cette chaîne ? Que peut-on prévoir comme évolution pour Y_n ?
3. On note Π la matrice de transition de cette chaîne. Montrer que :

$$\mathbb{E}(Y_n|Y_0 = j) = \sum_{k=0}^{2N} k(\Pi^n)(j,k)$$

où $(\Pi^n)(j,k)$ est la $j^{\text{ème}}$ ligne, $k^{\text{ème}}$ colonne de la matrice Π^n.
En admettant la propriété de la martingale $\mathbb{E}(Y_n|Y_{n-m}) = Y_{n-m}$ en déduire $\lim_{n\to+\infty} \Pi^n$.

4. En fonction de la loi initiale donner les probabilités d'absorption par 0 et $2N$.

Application à une loi initiale déterministe • $\mathcal{F}_0$ contient le même nombre de gènes A que de gènes a

- $\mathcal{F}_0$ contient trois fois plus de gènes A que de gènes a.
- $\mathcal{F}_0$ contient trois fois moins de gènes A que de gènes a.

Application à une loi initiale quelconque
- $\mathcal{F}_0$ a autant de chances de contenir $0, 1, 2, \ldots, 2N$ gènes A.
- $\mathcal{F}_0$ a deux chances sur trois de contenir le même nombre de gènes A que de gènes a et une chance sur trois de contenir deux fois plus de gènes A que de gènes a.
- Le nombre de gènes A dans la population $\mathcal{F}_0$ suit une loi binomiale $\mathcal{B}(2N, p)$.

5. Reprendre le problème avec un avantage sélectif du gène A :

$$\mathbb{P}(Y_{n+1} = k | Y_n = j) = C_{2N}^k (p_j)^k (1 - p_j)^{2N-k}$$

avec $p_j = \frac{1-e^{-js/N}}{1-e^{-2s}}$ où $s > 0$.
On admettra que e^{-2sY_n} a la propriété de la martingale, c'est-à-dire que $\mathbb{E}\left(e^{-2sY_n} | Y_{n-m}\right) = \mathbb{E}\left(e^{-2sY_{n-m}}\right)$

4.5 *Chaîne d'Ehrenfest*

On étudie l'échange des molécules entre deux compartiments notés 1 et 2. On a d molécules numérotées $1, 2 \ldots, d$. À chaque instant, on tire au hasard une molécule que l'on change de compartiment.
X_n est le nombre de molécules dans le premier compartiment à l'instant n.

1. Montrer que (X_n) est une chaîne de Markov dont on donnera la matrice de transition Π.
2. Étudier la nature des états de cette chaîne. Que peut-on en déduire ?
3. En supposant X_0 de loi binomiale $\mathcal{B}(d, 1/2)$ déterminer la loi de X_1. Qu'en déduit-on ?
4. De la distribution invariante et de la période de la chaîne, déduire $\lim_{n\to+\infty} \Pi^n$. Conclure sur la limite de la chaîne.
5. Application : décrire ce qui se passe si $d = 3$ et si :
   (a) au départ,toutes les molécules sont dans le premier compartiment.
   (b) au départ, chaque molécule est placée au hasard.
6. *Chaîne d'Ehrenfest modifiée*
   Lorsque la molécule a été extraite d'un compartiment, on tire au sort (avec des probabilités égales) dans quel compartiment on la met.
   Déterminer la matrice de transition et la distribution invariante de la chaîne ainsi modifiée. Étudier la loi limite.

Chapitre 5

Processus de branchement (ou de ramification ou de Galton-Watson)

1 Description du proccessus

Ces processus sont utilisés pour décrire l'évolution des populations, la croissance (ou la décroissance) de leurs effectifs, les probabilités d'extinction, etc.
On considère qu'à la génération 0 on a 1 individu. Cet individu peut avoir des descendants qui constituent la génération 1. Chaque individu peut avoir des descendants qui constituent la génération 2.
Les exemples d'application de ces processus sont nombreux en physique, en épidémiologie, en généalogie ou encore en génétique.
La survivance des noms de famille est un des premiers exemples de ce processus. Sir Galton (fondateur de l'eugénisme et cousin de Darwin) posa le problème de l'extinction des noms de famille au cours des générations. Watson fut le premier à proposer une solution mathématique à ce problème. Dans ces modèles, les seuls descendants considérés sont les enfants mâles.
On note Y_n l'effectif à la n$^{\text{ème}}$ génération et $X_{j,n}$ le nombre de descendants du j$^{\text{ème}}$ individu à la n$^{\text{ème}}$ génération ($j = 1, \ldots, Y_n$).
Par définition, une génération est égale à la réunion des descendants de tous les individus de la génération précédente :

$$Y_n = X_{1,n-1} + \cdots + X_{Y_{n-1},n-1} = \sum_{j=1}^{Y_{n-1}} X_{j,n-1}$$

On suppose que les individus se reproduisent indépendamment les uns des autres et que le nombre de descendants suit une loi qui ne dépend ni de l'individu parent ni de la génération.

$$\{X_{n,j} : n \geq 0 \text{ et } 1 \leq j \leq Y_{n-1}\} \text{ i.i.d.}$$

Le phénomène suit une loi stable au cours du temps et indépendante des individus. On note

$$p_k = \mathbb{P}(X_{n,j} = k) = \mathbb{P}(X = k)$$

La loi de X est entièrement déterminée par les p_k (qui ne dépendent ni de n ni de j).
On note que si il existe k tel que $\mathbb{P}(X = k) = 1$ (*i.e.* on est sûr que chaque individu a exactement k descendants) alors le processus est déterministe.

Théorème 5.1 *Si $\forall k$ $\mathbb{P}(X = k) < 1$ et si $\mathbb{P}(X \leq 1) = 1$ alors $\mathbb{P}(Y_n = 0) \stackrel{n\to+\infty}{\longrightarrow} 1$. C'est-à-dire que la population s'éteint.*

Théorème 5.2 *Si $\forall k$ $\mathbb{P}(X = k) < 1$ et si $\mathbb{P}(X \leq 1) < 1$ alors*

$$\forall k \neq 0 \quad \mathbb{P}(Y_n = k) \stackrel{n\to+\infty}{\longrightarrow} 0$$

et

$$\mathbb{P}(Y_n = 0) \stackrel{n\to+\infty}{\longrightarrow} \pi$$

Soit la population s'éteint (avec une probabilité π), soit elle explose (avec une probabilité $1-\pi$)

Le but est donc de déterminer la probabilité d'extinction π.

1.1 Fonction génératrice

Définition 5.1 *Si X est une variable aléatoire à valeurs dans $\mathbb{N}$, on définit :*

$$\forall s \in [0,1] \quad G(s) = \mathbb{E}(s^X) = \sum_{k=0}^{+\infty} s^k \mathbb{P}(X = k)$$

On note que :
- $G(0) = \mathbb{P}(X = 0)$
- $G(1) = 1$
- $G'(1) = \mathbb{E}(X)$

Exemple : Si $X \sim \mathcal{P}(\lambda)$ alors $G(s) = \sum_{k=0}^{+\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} e^{s\lambda} = e^{\lambda(s-1)}$

1.2 Étude du processus de branchement (Y_n)

On note G_n la fonction génératrice de Y_n.
$G_n(0) = \mathbb{P}(Y_n = 0)$

Proposition 5.1 $G_n = G \circ G_{n-1}$ *où G est la fonction génératrice de la loi de la descendance X.*

Donc :

$$\underbrace{G_n(0)}_{\stackrel{n\to+\infty}{\longrightarrow}\pi} = G\left(\underbrace{G_{n-1}(0)}_{\stackrel{n\to+\infty}{\longrightarrow}\pi}\right) \qquad \Longleftrightarrow \qquad \pi = G(\pi)$$

Théorème 5.3 *π est la plus petite racine positive de l'équation $G(s) = s$*

Théorème 5.4 *Il y a deux cas :*
- *Si $G'(1) = \mathbb{E}(X) \leq 1$ alors $\pi = 1$*
- *Si $G'(1) = \mathbb{E}(X) > 1$ alors $\pi < 1$ et π est la plus petite racine positive de $G(s) = s$*

2 Exercices

1. Étudier les processus de branchement dont la loi de la descendance est définie par :
   (a) $\mathbb{P}(X = 0) = \mathbb{P}(X = 3) = 1/2$
   (b) $\mathbb{P}(X = 0) = p_0$, $\mathbb{P}(X = 1) = 1 - p_0$. Dans ce cas étudier la loi de la variable aléatoire T : instant où la population s'éteint.
   (c) $\mathbb{P}(X = 0) = p$, $\mathbb{P}(X = 2) = 1 - p$.
   (d) $\mathbb{P}(X = 0) = p$, $\mathbb{P}(X = N) = 1 - p$.
   (e) X suit une loi binomiale $\mathcal{B}(N, p)$
   (f) $X = Y - 1$ où Y suit une loi géométrique de paramètre p : $\mathbb{P}(Y = k) = p(1 - p)^{k-1}$, $k \in \mathbb{N}$
   (g) X suit une loi de Poisson de paramètre λ.
2. On suppose que tout homme dans une certaine société a exactement trois enfants, avec des probabilités ½ pour chaque enfant d'être un garçon ou une fille.
   (a) Étudier le processus de branchement associé et la probabilité d'extinction du nom de famille d'un homme donné.
   (b) Si un homme donné a deux garçons et une fille, quelle est la probabilité que son nom ne s'éteigne jamais ?
3. Étudier le processus de branchement pour lequel $Y_0 = N$.
4. Au temps 0, une culture sur le sang commence avec une cellule rouge. Au bout d'une minute, la cellule rouge meurt et donne naissance à l'une des combinaisons suivantes (avec les probabilités indiquées) :
   - 2 cellules rouges avec une probabilité de ¼
   - 1 cellule rouge, 1 cellule blanche avec une probabilité de ⅔
   - 2 cellules blanches avec une probabilité de 1/12

   Chaque cellule rouge fonctionne de la même façon et chaque cellule blanche meurt au bout d'une minute sans reproduction.
   (a) Modéliser l'évolution de la population
   (b) Quelle est la probabilité qu'aucune cellule blanche n'apparaisse avant le temps $n + 1/2$?
   (c) Quelle est la probabilité que la culture toute entière meure ?