

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE
UNIVERSITE MOULOU MAMMERI DE TIZI-OUZOU
FACULTE DE SCIENCES - DEPARTEMENT DE MATHEMATIQUES

MEMOIRE DE MASTER DE
MATHEMATIQUES APPLIQUEES
Option : Processus Aléatoires et
Statistique de la Décision

*Chaînes de Markov et Algorithmes de
simulation*

présenté le 09 /11 /2014

par Mlle BENAKKOUCHE AMAL

sous la Direction de

Mr BOUDIBA MOHAND AREZKI

devant le Jury :

Mr HAMADOUCHE DJAMEL, Professeur, UMMTO, Président
Mr BOUDIBA MOHAND AREZKI, MC, UMMTO, Rapporteur
Mr MAMOU MOHAMED, Chargé de Recherches, UMMTO, Examineur

Remerciements

J'adresse mes vifs remerciements à Monsieur Hamadouche Djamel, pour l'honneur qu'il me fait en présidant le jury.

Je remercie également Monsieur Mamou Mohamed, pour avoir accepté de faire partie du jury.

Je présente mes sincères remerciements à mon directeur de mémoire Monsieur Boudiba Mohand Arezki, pour m'avoir proposer ce thème, pour, ses orientations, son suivi ses critiques et ses conseils.

Enfin, il m' est agréable d'adresser mes remerciements à tous ceux qui de près ou de loin ont contribué à la réalisation de ce travail.

Table des matières

1	Chaîne de Markov Marches aléatoires	7
1.1	Chaînes de Markov : Définition et propriétés générales	7
1.2	Marche aléatoire et simulation de trajectoires	11
2	Algorithmes de simulation	25
2.1	Simulation d'une loi de probabilités	25
2.2	Simulation de chaînes de Markov	29
2.3	Chaîne de Metropolis	36
3	Echantillonnage de Gibbs et Applications à la Physique	39
3.1	Introduction	39
3.2	Introduction au modèle d'Ising	43

Introduction

Nous nous intéressons aux algorithmes de simulation et à leur applications. Notre but est de donner les justifications mathématiques aux différents algorithmes utilisés. Souvent l'utilisateur est réduit à l'utilisation de techniques sans explication quant à leurs fondements. Il y a beaucoup d'éléments à clarifier compte tenu d'un développement sans précédent des techniques de simulation. Il nous faut souligner que les méthodes et techniques de simulation sont un terrain idéal pour voir l'impact profond du Calcul de Probabilités et de ses outils mathématiques dans la vie scientifique.

Il y a dans la littérature une foison de travaux consacrés à la simulation, qui sortent régulièrement. Quelques uns ont retenu notre attention. En particulier les travaux de Saloff Coste et Diaconis, Häggström, Ycart, Benaim, Levin,... (Cf. [1], [2], [3], [4], [7], [8],...) Nous en faisons la synthèse de quelques éléments de base, dans ce mémoire.

Concrètement, au chapitre I nous présentons une synthèse sur les chaînes de Markov et leurs propriétés les plus importantes. Nous donnons l'exemple des marches aléatoires. Dans le cas des marches aléatoires sur \mathbb{Z} nous faisons leur caractérisation et illustrons le propos par quelques simulations de trajectoires.

Au chapitre II, nous présentons quelques algorithmes de simulation simples (simulation d'une loi de probabilités) et nous synthétisons une méthode de simulation de chaînes de Markov.

Nous présentons enfin l'algorithme de Metropolis et l'algorithme de Gibbs (Chap III) de simulation d'une loi de probabilités par chaînes de Markov, désigné dans la littérature par la méthode MCMC et enfin (Chap III) une

introduction au modèle d'Ising .

Nous terminons notre exposé par une conclusion sur les perspectives pour développer ce travail dans l'avenir.

Chapitre 1

Chaîne de Markov Marches aléatoires

1.1 Chaînes de Markov : Définition et propriétés générales

Définition 1. Soit (E, \mathcal{E}) un espace mesurable et $(X_n)_{n \geq 0}$ une suite de variables aléatoires à valeurs dans E . On dit que $(X_n)_{n \geq 0}$ est une chaîne de Markov d'espace des états E , si $\forall B \in \mathcal{E}$

$$P[X_{n+1} \in B | X_n, X_{n-1}, \dots, X_0] = P[X_{n+1} \in B | X_n]$$

Remarque 1. Dans le cas où E est un ensemble fini ou dénombrable et $(X_n)_{n \geq 0}$ une suite de variables aléatoire à valeurs dans E , $(X_n)_{n \geq 0}$ est une chaîne de Markov d'espace des états E et de matrice de transition $P = (p_{x,y})_{x,y \in E}$ si :

$$\begin{aligned} & \forall x_0, x_1, \dots, x_n, x, y \in E, \text{ et } \forall n \in \mathbb{N} \\ & P[X_{n+1} = y | X_0 = x_0, X_1 = x_1, \dots, X_n = x] = \\ & = P[X_{n+1} = y | X_n = x] = p_{x,y} \end{aligned}$$

Si les nombres $p_{x,y}$ ne dependent pas de n , alors la chaîne de Markov est dite homogène

N.B Dans toute la suite, les chaînes de Markov considérées sont toutes homogènes et à espace des états au plus dénombrable. Soit $E = \{x, y, \dots\}$ un ensemble fini $(X_n)_n$ une chaîne de Markov sur E de matrice de transition $P = (p_{x,y})_{x,y \in E}$

Proposition 1. Pour tout entier n, m , et pour i et $j \in E$ soit

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i) = P(X_{m+n} = j | X_m = i)$$

C'est la probabilité de transition de l'état i à l'état j , en n pas. Si P est la matrice de transition, alors on a

$$(p_{ij}^{(n)})_{i,j \in E} = P^n$$

N.B Les probabilités de transition d'une chaîne de Markov à temps discret vérifient la relation fondamentale suivante (appelée equation de Chapman-Kolmogorov)

$$p_{i,j}^{(n+m)} = \sum_{z \in E} p_{x,z}^{(n)} p_{z,y}^{(m)}$$

pour tous instants $\forall n, m > 0, x, y \in E$. Notons que cette équation n'est pas une condition suffisante pour qu'un processus soit une chaîne de Markov

$$\begin{aligned} p_{x,y}^{(n+m)} &= P(X_{n+m} = y | X_0 = x) = \sum_{z \in E} P(X_{n+m} = y, X_m = z | X_0 = x) \\ &= P(X_m = z | X_0 = x) P(X_{n+m} = y | X_0 = x, X_m = z) \\ &= P(X_{n+m} = y | X_m = z) P(X_m = z | X_0 = x) \\ &= \sum_{z \in E} p_{zy}^{(n)} p_{xz}^{(m)} = \sum_{z \in E} p_{xz}^{(m)} p_{zy}^{(n)} \end{aligned}$$

Définition 2. Soit $\pi = (\pi_x)_{x \in E}$ une famille de nombres réels positifs tel que $\sum_x \pi_x = 1$

1. π est une loi stationnaire (invariante) pour la chaîne de Markov $(X_n)_n$ si pour $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ alors si $\pi = \pi P$, c'est à dire, si pour tout $y \in E$ on a :

$$\pi_y = \sum_x \pi_x p_{xy}$$

2. La loi π est dite réversible ou que la matrice P est π réversible si

$$\forall x, y \in E, \pi_x p_{x,y} = \pi_y p_{y,x}$$

N.B

$$\pi_y = \sum_x \pi_x p_{x,y} \Leftrightarrow \pi_y = \sum_x \pi_x p_{x,y}^{(n)} \quad \forall n$$

Démonstration. \Leftarrow

n=1 fait

donc on a

$$\begin{aligned} \pi_y &= \sum_x \pi_x p_{xy}^{(1)} = \sum_x \pi_x p_{xy} \\ &\Rightarrow \end{aligned}$$

On a $\pi_y = \sum_x \pi_x p_{xy} \quad \forall x, y$

donc $\pi_x = \sum_z \pi_z p_{zx}$

On remplace dans (1)

$$\pi_y = \sum_x \left(\sum_z \pi_z p_{zx} \right) p_{xy} = \sum_z \pi_z \sum_x p_{zx} p_{xy}$$

D'après la propriété de Kolmogorov $p_{zy}^{(2)} = \sum_x p_{zx} p_{xy}$

On a $\pi_y = \sum_z \pi_z p_{zy}^{(2)}$ Supposons que c'est vrai à l'ordre n i.e

$$\pi_y = \sum_z \pi_z p_{zy}^{(n)}$$

On doit montrer que c'est vrai à l'ordre $n+1$

$$\pi_y = \sum_z \pi_z p_{zy}^n$$

mais $\pi_z = \sum_x \pi_x p_{xz}$

donc

$$\pi_y = \sum_z \left(\sum_x \pi_x p_{xz} \right) p_{zy}^{(n)} = \sum_x \pi_x \sum_z p_{xz} p_{zy}^{(n)} = \sum_x \pi_x p_{xy}^{(n+1)}$$

Donc

$$\pi_y = \sum_x \pi_x p_{x,y} \Leftrightarrow \pi_y = \sum_x \pi_x p_{x,y}^{(n)} \quad \forall n$$

□

Remarque 2. Si π est réversible, alors elle est stationnaire. En effet ;

$$\pi_x p_{xy} = \pi_y p_{yx}$$

En sommant par rapport à x , on obtient :

$$\sum_{x \in E} \pi_x p_{xy} = \sum_{x \in E} \pi_y p_{yx} = \pi_y \sum_{x \in E} p_{yx} = \pi_y, y \in E$$

car $\sum_x p_{yx} = 1$ C'est à dire $\forall y \in E, \pi_y = \sum_x \pi_x p_{xy}$

Définition 3. Une chaîne de Markov $(X_n)_n$, est dit irréductible si elle ne possède qu'une unique classe de communication, c'est à dire que tous ses éléments communiquent ($x \longleftrightarrow y$).

Définition 4. Soit $(X_n)_n$ une chaîne de Markov irréductible. La période d'un état $x \in E$ est le PGCD (plus grand commun diviseur) de $n \in \mathbb{N}^*$; $P_{xx}^n > 0$ un état est dit apériodique si sa période est 1. Si non il est dit périodique.

Note : la périodicité est une propriété de classe. Si une classe de communication est de période 1, on dit qu'elle est apériodique. une chaîne de Markov irréductible est dite apériodique si elle admet une classe apériodique

Récurrence Soit $(X_n)_n$ une chaîne de Markov d'espace des états E et de matrice de transition $P = (p_{xy})_{x,y \in E}$ et pour $y \in E$ soit T_y la v.a à valeurs dans \mathbb{N} définie par

$$T_y = \begin{cases} \inf \{n > 0, X_n = y\}, & \text{si } \bigcup_n \{X_n = y\} \neq \emptyset \\ \infty, & \text{sinon} \end{cases}$$

Intuitivement, T_y est le premier instant de visite de y . Si $X_0 = y$, T_y est le premier instant de retour à y .

Définition 5. Soit $(X_n)_n$ une chaîne de Markov sur E de matrice de transition $P = (p_{xy})_{x,y}$. Pour $y \in E$ soit T_y le premier instant de visite à y . L'état y est recurrent si $P[T_y < \infty | X_0 = y] = 1$ et transient sinon.

Autrement dit, si

$$f_{xy}^{(n)} = P[T_y = n | X_0 = x] \text{ et si } f_{xy}^* = \sum_n f_{xy}^{(n)},$$

l'état y est récurrent si $f_{yy}^* = 1$ et transient si $f_{yy}^* < 1$

Proposition 2. Si $(X_n)_n$ est une chaîne de Markov sur E de matrice de transition $P = (p_{xy})$, $p_{xy}^{(n)} = P[X_n = y | X_0 = x]$

$x \in E$, alors

$\sum p_{xx}^{(n)} = \infty$ x est récurrent

$\sum p_{xx}^{(n)} < \infty$ x est transient

Proposition 3. Soit $(X_n)_n$ une chaîne de Markov sur E de matrice de transition $P = (p_{xy})$, si T_x Pour $x \in E$, alors T_x le premier instant de retour à x et on suppose que $(X_n)_n$ admet la loi stationnaire $\pi = (\pi_x)_{x \in E}$. Alors on a

$$\mathbb{E}[T_x] = \frac{1}{\pi_x}$$

Proposition 4. (1). Si $y \in E$, C_y est la classe de y alors si y est récurrent les états de C_y sont récurrents. idem pour la transience.

(2). Si $q_{yy} = P[X_k = y, \text{ une } \infty \text{ de fois} | X_0 = y]$ alors y est récurrent $\Leftrightarrow q_{yy} = 1$.
Si y est transient $\Leftrightarrow q_{yy} = 0$

1.2 Marche aléatoire et simulation de trajectoires

Marche aléatoire sur \mathbb{Z} Soit $(Y_n)_n$ une suite de variables aléatoires *iid* tel que $P(Y_n = 1) = p$ et $P(Y_n = -1) = q$; $p + q = 1$ et X_0 une variable aléatoire indépendante de $(Y_n)_n$. On considère la suite $(X_n)_n$ de variables aléatoires définie par

$$X_0 \text{ et pour } n > 0, X_n = \sum_{k=1}^n Y_k.$$

Alors $(X_n)_n$ est une chaîne de Markov sur \mathbb{Z} de probabilités de transitions les nombres $p_{x,y}$ définis pour x et $y \in \mathbb{Z}$ par

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

Démonstration. Nous avons, pour $x, y, \dots, x_0 \in \mathbb{Z}$,

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

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

Y_{n+1} indépendant de X_n, X_{n-1} donc

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

D'autre part

$$P(X_{n+1} = y / X_n = x) = \frac{P(Y_{n+1} = y - x, X_n = x)}{P(X_n = x)} = P(Y_{n+1} = y - x) \dots \dots \dots (2)$$

de (1) et (2) la chaîne $(X_n)_n$ est une chaîne de Markov \square

Nous donnons ci-dessous quelques simulations de trajectoires suivant les valeurs de p , pour la marche aléatoire sur \mathbb{Z} .

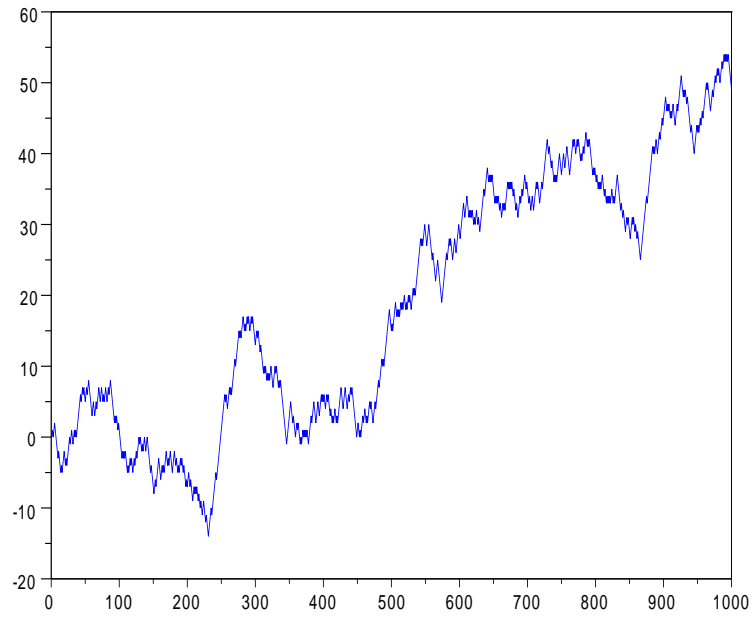


FIGURE 1.1 – (a) Marche aléatoire sur \mathbb{Z} avec $p = q$

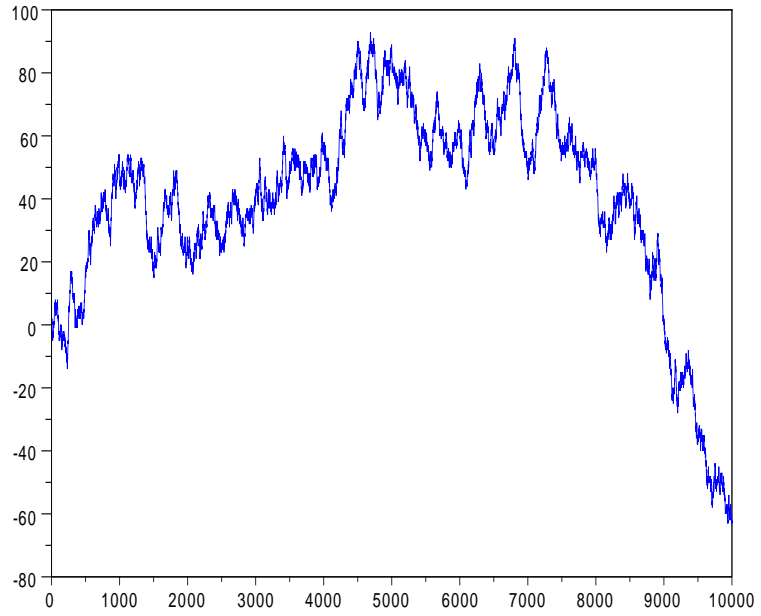


FIGURE 1.2 – (b) Marche aléatoire sur \mathbb{Z} avec $p = q$

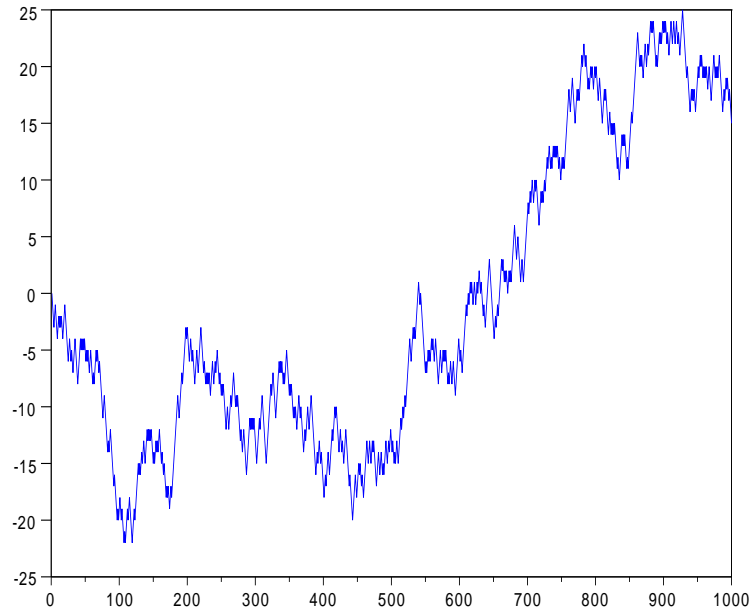


FIGURE 1.3 – (c) Marche aléatoire sur \mathbb{Z} avec $p = 0.48$ et $q = 0.52$

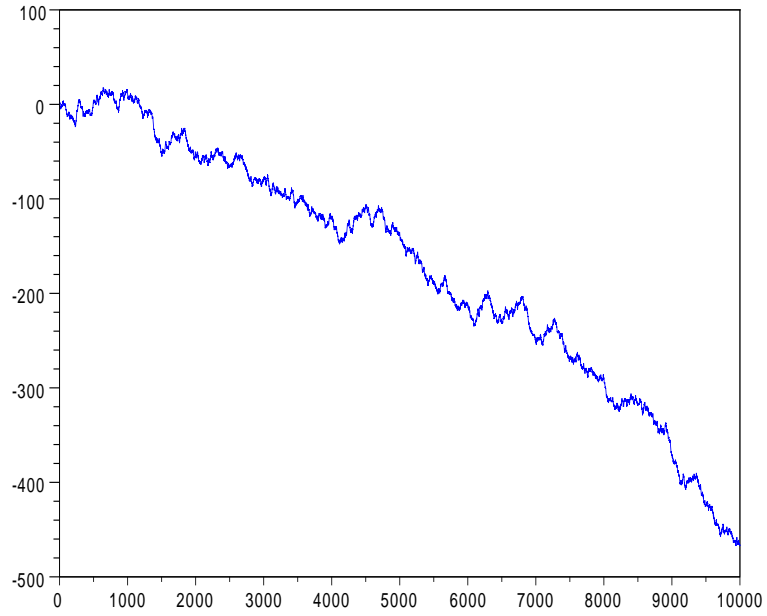


FIGURE 1.4 – (d) Marche aléatoire sur \mathbb{Z} avec $p = 0.48$ et $q = 0.52$

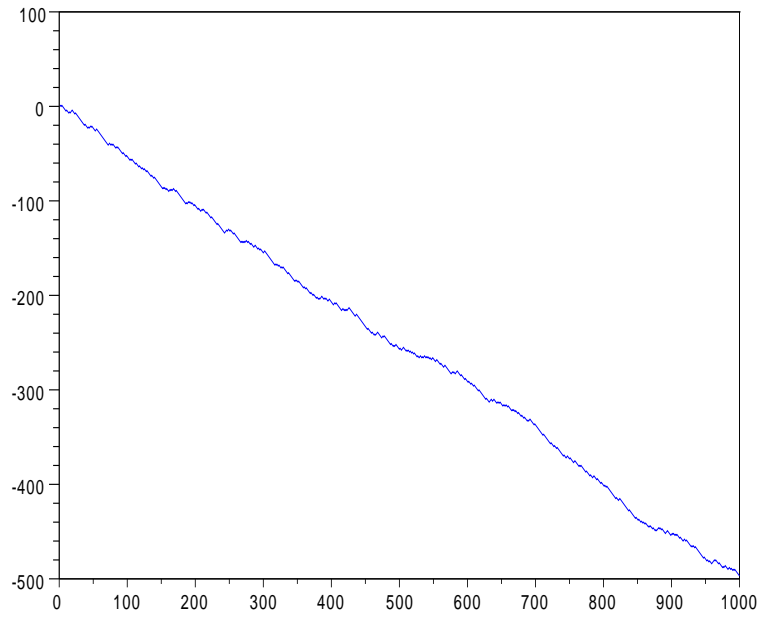


FIGURE 1.5 – (e) Marche aléatoire sur \mathbb{Z} avec $p = 0.25$ et $q = 0.75$

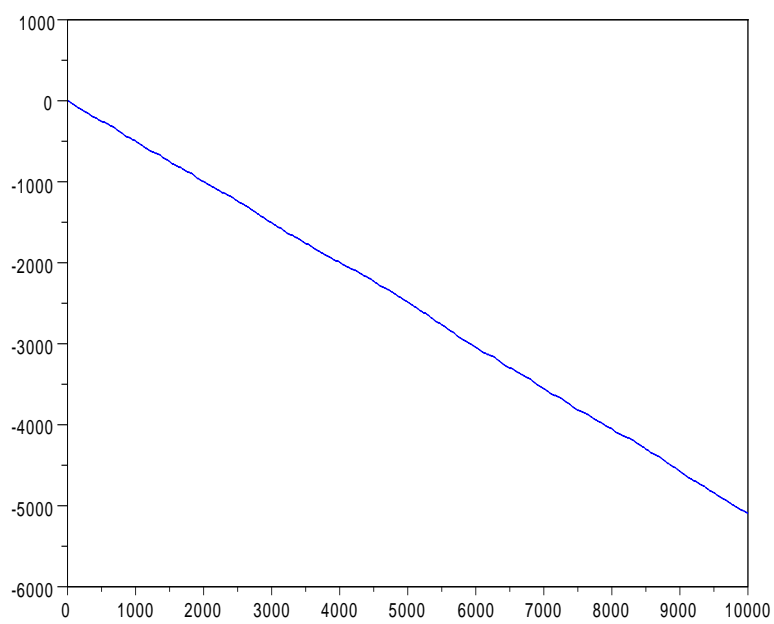


FIGURE 1.6 – (f) Marche aléatoire sur \mathbb{Z} avec $p = 0.25$ et $q = 0.75$

Synthèse des observations (1) Dans le cas où $p = q$, les figures (2.1) et (2.2) montrant que la marche aléatoire repasse en 0 même si le nombre d'itérations est faible relativement.

(2) Si $p \neq q$ lorsque p est proche de q (fig 2.3) la marche aléatoire repasse en 0 au admet des itérations comme dans le premier cas mais elle ne repasse plus (fig 2.4).

(3) Si $p \neq q$ et p sensiblement \neq de q , la marche aléatoire s'échappe vers $-\infty$ (fig 2.5, 2.6)

Conclusion

Cela illustre bien l'étude théorique, car nous savons que la marche aléatoire symétrique sur \mathbb{Z} est récurrente pour $p = q$ et transiente sinon. En effet nous avons la proposition suivante

Proposition 5. Soit $(X_n)_n$ la marche aléatoire sur \mathbb{Z} engendrée par la suite $(Y_n)_n$ de v.a tel que $P(Y_n = 1) = p$ et $P(y_n = -1) = q$, $p + q = 1$. Si $p_{x,y}^n = P(X_{n+1} = y | X_n = x)$. Alors on a si

$$p = q, \sum_n p_{xx}^{(n)} = \infty$$

$$p \neq q, \sum_n p_{xx}^{(n)} < \infty$$

On a si $p = q$, $(X_n)_n$ est récurrente
Si $p \neq q$, $(X_n)_n$ transiente.

Démonstration. pour $X_0 = 0$ $p \neq q$: on utilise le lemme de Borel-Cantelli ie on montre que $q_{00} = 0$ est 0 transient $\Rightarrow (X_n)_n$ transient

$$q_{0,0} = P[X_n = 0, \text{ une } \infty \text{ de fois}] = 0$$

une méthode est l'application du Lemme de Borel-Cantelli. En effet si nous montrons que

$$\sum_n P[X_n = 0] < \infty,$$

la conclusion est immédiate. Pour cela calculons $P[X_n = 0]$. Remarquons que l'événement $\{X_n = 0\}$ ne se réalise que si n est paire. Comme

$$\{X_{2n} = 0\} = \{Y_1 + Y_2 + \dots + Y_{2n} = 0\}$$

il s'en suit que l'événement $\{X_n = 0\}$ ne se réalise que si les Y_n prennent n fois la valeur $+1$ et n fois la valeur -1 (on ne revient en 0 au bout de $2n$ pas que si on fait n pas à droite et n pas à gauche, au total). Les pas sont indépendants et chaque pas à droite est fait avec la probabilité p et chaque pas à gauche avec la probabilité q . Le nombre total de possibilités en $2n$ pas est le nombre de façons de choisir n pas à gauche parmi $2n$ pas, soit C_{2n}^n . Par suite

$$P[X_n = 0] = C_{2n}^n p^n q^n$$

et la série

$$\sum_n P[X_n = 0] = \sum_n P[X_{2n} = 0] \sum_n C_{2n}^n p^n q^n$$

est convergente. En effet, appliquons la règle de D'Alembert, on a

$$\frac{C_{2n+2}^{n+1} p^{n+1} q^{n+1}}{C_{2n}^n p^n q^n} = pq \frac{C_{2n+2}^{n+1}}{C_{2n}^n} = pq \frac{4n^2 + 6n + 2}{n^2 + 2n + 1} \rightarrow 4pq$$

quand $n \rightarrow \infty$. Comme pour $p \neq q$ on a $4pq = 4p(1-p) < 1$, il s'en suit que la série est convergente d'après la règle de D'Alembert et par suite, d'après le lemme de Borel-Cantelli,

$$P[\limsup_n \{X_n = 0\}] = 0$$

i.e.

$$P[X_n = 0, \text{ une } \infty \text{ de fois}] = 0$$

On en déduit que $q_{00} = 0$ et donc que 0 est transient et comme la transience est une propriété de classe tous les états sont donc transients et part suite, la chaîne est transiente.

pour $p = q$

Dans ce cas le lemme de Borel Cantelli ne s'applique pas, on applique la proposition 2.

Nous avons :

$$p = q = \frac{1}{2} \text{ et } \sum_n p_{00}^{(n)} = \sum_n C_{2n}^n \frac{1}{2^{2n}}.$$

Quand $n \rightarrow \infty$, nous avons (en tenant compte que $n! \sim e^{-n}n^{n+\frac{1}{2}}\sqrt{2\pi}$),

$$C_{2^n}^n \frac{1}{2^{2n}} = \frac{(2n)!}{(n!)^2} \frac{1}{2^{2n}} \sim \sqrt{2} \frac{1}{\sqrt{n}}.$$

La série

$$\sum_n \frac{1}{\sqrt{n}}$$

est une série divergente. Il en résulte que la série

$$\sum_n p_{00}^{(n)}$$

est divergente et 0 est récurrent. Donc tous les états sont récurrents car la récurrence est une propriété de classe et la chaîne est récurrente.

□

Marche aléatoire sur \mathbb{R}

Soit X_0 une variable aléatoire à valeurs dans \mathbb{R} , $(Y_n)_n$ une suite de variables aléatoires à valeurs dans \mathbb{R} indépendantes de X_0 et de même loi μ , tel que si $A \in \mathcal{B}_{\mathbb{R}}$, $\mu(A) = \mathbb{P}[Y_1 \in A]$ et $\forall n > 0$, $X_n = \sum_{k=1}^n Y_k$, alors $(X_n)_n$ est une chaîne de Markov sur \mathbb{R} , de noyau de transition \mathbb{P} défini pour tout $x \in \mathbb{R}$ et $A \in \mathcal{B}_{\mathbb{R}}$ par

$$\mathbb{P}(A, x) = \mu(A - x)$$

Démonstration. Nous avons :

$$\forall A \in \mathcal{B}_{\mathbb{R}}, \forall n, P[X_{n+1} \in A | X_n, X_{n-1}, \dots, X_0] = P[X_{n+1} \in A | X_n]$$

En effet, par définition de la probabilité conditionnelle

$$\forall A_n, A_{n-1}, \dots, A_0 \quad P[X_{n+1} \in A | X_n, X_{n-1}, \dots, X_0]$$

vérifie

$$\begin{aligned} P[X_{n+1} \in A, X_n \in A_n, \dots, X_0 \in A_0] &= \int_{\{X_n \in A_n, \dots, X_0 \in A_0\}} P[X_{n+1} \in A | X_n, \dots, X_0] dP \\ &= \int_{\{X_n \in A_n, \dots, X_0 \in A_0\}} \mathbb{I}_A(X_{n+1}) dP \\ &= \int_{\{X_n \in A_n, \dots, X_0 \in A_0\}} \mu(A - X_n) dP. \end{aligned} \quad (1.1)$$

Soit $\phi(X_n, X_{n-1}, \dots, X_0)$ une fonction mesurable de \mathbb{R} dans $[0, 1]$.
Donc si

$$\phi(X_n, X_{n-1}, \dots, X_0) = P[X_{n+1} \in A \mid X_n, \dots, X_0]$$

Alors

$$\phi(X_n, X_{n-1}, \dots, X_0) = \mu(A - X_n)\text{-p.s.}$$

Par unicité de la probabilité conditionnelle.

$$P[X_{n+1} \in A \mid X_n, X_{n-1}, \dots, X_0] = P[X_{n+1} \in A \mid X_n]$$

• Pour caractériser $(X_n)_n$, on a besoin de préciser le noyau de transition associé \mathbb{P} qu'on définit par :

$$\begin{aligned} \mathbb{P}(A, x) &= P[Y_{n+1} \in A - x] \\ &= \mu(A - x) \end{aligned} \tag{1.2}$$

$$\tag{1.3}$$

Par définition de la probabilité conditionnelle nous avons :

$$\forall A \in \mathcal{B}_R, p[X_{n+1} \in A \mid X_n]$$

est l'unique variable aléatoire $\sigma(X_n)$ -mesurable qui vérifie :

$$\forall B \in \mathcal{B}_R, P[X_{n+1} \in A, X_n \in B] = \int_{\{X_n \in B\}} P[X_{n+1} \in A \mid X_n] dP$$

Cette variable aléatoire étant $\sigma(X_n)$ -mesurable, d'après (Breiman) Il existe une fonction $\varphi : \mathbb{R} \rightarrow [0, 1]$ mesurable

Tel que

$$P[X_{n+1} \in A \mid X_n] = \varphi(X_n).$$

Donc, avec cette notation pour A et $B \in \mathcal{B}_R$:

$$\begin{aligned}
P[X_{n+1} \in A, X_n \in B] &= \int_{\{X_n \in B\}} \varphi(X_n) dP \\
&= \int_{\{X_n \in B\}} \mathbb{E}[\mathbb{I}_A(X_{n+1}) | X_n] dP \\
&= \int_{\{X_n \in B\}} \mathbb{I}_A(X_{n+1}) dP \\
&= \int_{\{X_n \in B\}} \mathbb{I}_A(Y_{n+1} + X_n) dP \tag{1.4}
\end{aligned}$$

Donc

$$\begin{aligned}
\int_{\{X_n \in B\}} \varphi(X_n) dP &= \int_{\{X_n \in B\}} \mathbb{E}[\mathbb{I}_A(X_{n+1}) | X_n] dP \\
&= \int_{\{X_n \in B\}} \mathbb{I}_A(X_{n+1}) dP \\
&= \int_{\{X_n \in B\}} \mathbb{I}_A(X_n + Y_{n+1}) dP \\
&= \int_B \int \mathbb{I}_A(x + y) P[X_n \in dx, y_{n+1} \in dy] \tag{1.5}
\end{aligned}$$

Par indépendance de Y_{n+1} et X_n

$$= \int_B \int \mathbb{I}_A(x + y) p[X_n \in dx] P[y_{n+1} \in dy] \tag{1.6}$$

En appliquant le théorème de Fubini

$$= \int_B P[X_n \in dx] \int \mathbb{I}_A(x + y) P[y_{n+1} \in dy] \tag{1.7}$$

Par unicité de l'espérance conditionnelle on a :

$$\varphi(X_n) = P[Y_{n+1} \in A - X_n]$$

Donc

$$\mathbb{P}(A, x) = P[Y_{n+1} \in A - x].$$

Alors $(X_n)_n$ est une chaîne de Markov sur \mathbb{R} de noyau de transition défini par $x \in \mathbb{R}$ et $A \in \mathcal{B}_R$ par $P(A, x) = \mu(A - x)$

□

Chapitre 2

Algorithmes de simulation

2.1 Simulation d'une loi de probabilités

1er cas : Cas de fonction de répartition inversible On s'intéresse aux algorithmes de simulation, c'est à dire à la construction d'échantillons d'une loi de probabilités donnée π . Si X est une variable aléatoire, simuler X , c'est construire un échantillon (X_1, \dots, X_n) de X . La proposition suivante contient un algorithme simple de simulation.

Proposition 1. *Soit X une v.a. discrète de f.r. F . Si F admet une fonction inverse F^{-1} , alors la variable aléatoire $F(X)$ est de loi uniforme sur $[0, 1]$.*

Démonstration.

Si $x \in [0, 1]$ alors on a

$$P[F(X) \leq x] = P[X \leq F^{-1}(x)] = F[F^{-1}(x)] = x$$

□

Application : Algorithme de simulation d'une loi inversible

Soit U la v.a. définie par $U = F(X)$. U uniforme sur $[0, 1]$ si (U_1, \dots, U_n) est un échantillon de U , alors l'échantillon (X_1, \dots, X_n) avec $X_i = F^{-1}(U_i)$ est un échantillon de X suivant la loi F . Cela suppose donc que nous disposons d'un échantillon (U_1, U_2, \dots, U_n) de la loi uniforme sur $[0, 1]$.

Les algorithmes implementés dans les générateurs de nombres aléatoires dans les machines, fournissent avec une bonne précision ces échantillons. Si

on veut une meilleure précision, on peut exploiter le théorème de Borel qui permet de construire des échantillons de la loi uniforme sur $[0, 1]$ avec une très bonne précision. Nous rappelons le théorème.

Théorème 1 (Théorème de Borel). *Soit X une variable aléatoire dans $[0, 1] \cap \mathbb{Q}^c$. Si $X = 0, X_1, X_2, \dots, X_n, \dots$ est le développement décimal de X alors les X_i sont uniformément réparties sur $\{0, 1, \dots, 9\}$ c'est à dire*

$$P[X_i = a] = \frac{1}{10}, \forall a \in \{0, 1, \dots, 9\}$$

Ce théorème contient un algorithme de génération de nombres aléatoires uniformément répartie sur $[0, 1]$. Le problème revient à disposer d'un algorithme de développement décimal d'un irrationnel de $[0, 1]$. Si on choisit π , on peut exploiter pour cela, la formule :

$$\frac{\pi}{4} = \sum_{i=0}^{\infty} (-1)^i \frac{1}{2i+1}.$$

Cette méthode nécessite cependant une puissance de calcul importante. A supposer que cela est le cas, nous avons alors l'algorithme suivant.

Algorithme de construction de nombres aléatoires uniformes sur $[0, 1]$

Comment à partir de ce théorème obtenir un échantillon (U_1, \dots, U_n) uniforme sur $[0, 1]$?

Si $X = 0, X_1, X_2$ ou X_1, X_2 sont uniformes sur $\{0, 1, \dots, 9\}$ et montrons que X est uniforme sur $[0, 1]$

$$X = \frac{1}{10}X_1 + \frac{1}{100}X_2$$

Si $x \in [0, 1]$, x de la forme $x = 0, x_1x_2$ Alors

$$P[X = x] = P[X_1 = x_1, X_2 = x_2] = \frac{1}{100}$$

2 cas Fonction de répartition quelconque Dans le cas d'une fonction de répartition quelconque, nous introduisons la notion d'inverse généralisé pour avoir un résultat du même type que dans le cas inversible.

Définition 1. Soit F une fonction de répartition. On appelle inverse généralisé de la fonction F , la fonction

$$\bar{F}^{-1} :]0, 1[\rightarrow \mathbb{R}$$

définie par

$$\bar{F}^{-1}(t) = \inf \{x \in \mathbb{R}, F(x) \geq t\}$$

Proposition 2. (Propriété de \bar{F}^{-1})

- 1) La fonction \bar{F}^{-1} est croissante sur $[0, 1]$
- 2) pour $t \in [0, 1]$ et $x \in \mathbb{R}$, on a l'équation : $\bar{F}^{-1}(t) \leq x \iff F(x) \geq t$
- 3) Si F est continue sur \mathbb{R} , pour tout $t \in [0, 1]$, on a $F(\bar{F}^{-1}(t)) = t$ i.e \bar{F}^{-1} est un inverse à droite de F
- 4) Si F est continue et strictement croissante sur \mathbb{R} , l'inverse généralisée coïncide avec l'inverse usuel.

Théorème 2. Soit F une fonction de répartition d'inverse généralisée \bar{F}^{-1}

Alors :

- 1) Si U est une variable aléatoire uniforme sur $[0, 1]$, la variable aléatoire $Z = \bar{F}^{-1}(U)$ admet une fonction de répartition égale à F
- 2) Si F est continue sur \mathbb{R} et si X est une variable aléatoire alors la variable $U = F(X)$ est de loi uniforme sur $[0, 1]$.

Remarque 1. Le théorème précédent montre qu'on peut en principe simuler n'importe quelle loi continue à partir de loi uniforme sur $[0, 1]$. Il n'a qu'un intérêt théorique car il n'est pas facile dans la pratique d'utiliser pour produire des échantillons observés de loi F car il faudrait connaître parfaitement la fonction \bar{F}^{-1} ce qui n'est pas le cas en général. Il ya pour les lois classiques des méthodes particulières efficaces.

3 cas Fonction de répartition continue

Simulation de la loi $N(0,1)$: Méthode de Box et Muller

La méthode est basée sur la proposition suivante :

Proposition 3. Soit U et V deux v.a indépendantes de lois uniformes, sur $[0, 1]$, et soit X et Y les v.a définies par $X = \sqrt{-2 \ln U} \cos 2\pi V$ et $Y = \sqrt{-2 \ln U} \sin 2\pi V$. alors X et Y sont deux v.a indépendantes et de lois $N(0, 1)$.

Application L'algorithme qui en découle et que si nous disposons d'un échantillon $(U_1, V_1, U_2, V_2, \dots, U_n, V_n)$, de taille donnée $2n$ de la loi uniforme, d'échantillon $(X_1, Y_1, X_2, Y_2, \dots, X_n, Y_n)$ est un échantillon de taille $2n$ de la loi $N(0, 1)$.

Démonstration. montrons que X et Y sont deux v.a indépendantes et de loi $N(0, 1)$.

Nous avons U et V uniformes sur $[0, 1]$ par hypothèse .

Comme

$$\begin{cases} X = \sqrt{-2 \ln U} \cos 2\pi V \dots\dots\dots(1) \\ Y = \sqrt{-2 \ln U} \sin 2\pi V \dots\dots\dots(2) \end{cases}$$

On a : de (1) et (2) :

$$\begin{cases} X^2 + Y^2 = -2 \ln U \\ \tan V = \frac{Y}{X} \end{cases}$$

Donc

$$\begin{cases} U = e^{-\frac{1}{2}(x^2+y^2)} \\ V = \frac{1}{2\pi} \arctan \frac{Y}{X} \end{cases}$$

(U, V) uniforme sur $[0, 1] \times [0, 1]$. Soit $f_{U,V}$ la densité du couple (U, V) . Nous avons donc

$$f_{U,V}(uv) = \begin{cases} 1 & \text{si } (u, v) \in [0, 1]^2 \\ 0 & \text{si non} \end{cases}$$

Si B borelien de $[0, 1]^2$

$$P[(U, V) \in B] = \int_B f_{U,V}(u, v) dudv \dots\dots\dots(I)$$

soit alors $(x, y) \rightarrow \phi(x, y) = (u, v)$ La formule de changement de variable pour les integrales doubles , appliquée à (I), donne alors

$$\int_B f_{U,V} dudv = \int_{\phi^{-1}(B)} f_{U,V} \circ \phi(x, y) |J_{\phi(x,y)}(x, y)| dx dy$$

$$J_{\phi} = \begin{pmatrix} \frac{du}{dx} & \frac{du}{dy} \\ \frac{dv}{dx} & \frac{dv}{dy} \end{pmatrix} = \begin{pmatrix} -xe^{-\frac{1}{2}(x^2+y^2)} & -ye^{-\frac{1}{2}(x^2+y^2)} \\ \frac{-y}{2\pi} \frac{1}{x^2+y^2} & \frac{x}{2\pi} \frac{1}{x^2+y^2} \end{pmatrix}$$

$$\frac{dv}{dx} = \frac{-y}{x^2} \frac{1}{2\pi} \frac{1}{\frac{x^2+y^2}{x^2}} = \frac{-y}{2\pi} \frac{1}{x^2 + y^2}$$

$$\frac{dv}{dy} = \frac{1}{2\pi} \frac{\frac{1}{x}}{x^2 + \frac{y^2}{x^2}} = \frac{x}{2\pi} \frac{1}{x^2 + y^2}$$

$$|\det J_\phi| = \left| -\frac{x^2 + y^2}{2\pi(x^2 + y^2)} e^{-\frac{1}{2}(x^2 + y^2)} \right|$$

Donc $J_\phi \neq 0$ il s'en suit que ϕ est bijective et on a :

$$\begin{aligned} (I) &= \int_{\phi^{-1}(B)} f_{U,V} \phi(x,y) \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(x^2 + y^2)} dx dy \\ &= \int_{\phi^{-1}(b)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dx dy \\ &= \int_{B_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \int_{B_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \end{aligned}$$

Si $\phi^{-1}(B) = B_1, B_2$ donc $P[(U, V) \in B] = P[(X, Y) \in \phi^{-1}(B)] = P[X \in B_1]P[Y \in B_2]$ et on observe que X et $Y \in N(0, 1)$ car les probabilités ci-dessus s'expriment par des densités de la loi normale $N(0, 1)$.

□

2.2 Simulation de chaînes de Markov

Supposons qu'on veut simuler une chaîne de Markov $(X_n)_{n \geq 0}$ définie sur ensemble fini $E = \{x, y, \dots\}$, de matrice de transition $P = (p_{x,y})_{x,y \in E}$ et de loi initiale $\mu^{(0)}$. On suppose donnée une suite U_0, U_1, \dots, U_n de variables aléatoires indépendantes identiquement distribuées (iid) tel que :

$$\forall i, U_i \sim U[0, 1]$$

Soit la fonction ψ (fonction d'initialisation) tel que $\psi : [0, 1] \rightarrow E$ définie par :

$$\psi(u) = \begin{cases} x & \text{si } u \in [0, \mu^{(0)}(x)[; \\ y & \text{si } u \in [\mu^{(0)}, \mu^{(0)}(x) + \mu^{(0)}(y)[\\ \vdots & \\ h & \text{si } u \in [\sum_{l=0}^{h-1} \mu^{(0)}(l), 1] \end{cases}$$

On a alors $X_0 = \psi(U_0)$.

Proposition 4. *la loi de $\psi(U_0)$ est égale à $\mu^{(0)}$*

Démonstration.

$$\begin{aligned}
P[\psi(U_0) = h] &= \int_0^1 \mathbb{I}_{\psi(u)=h} d\mu = \mu(\psi(u) = h) \\
&= \sum_{l=x}^h \mu^0(l) - \sum_{l=x}^{h-1} \mu^0(l) \\
&= \mu^0(h) = P[x_0 = h]
\end{aligned}$$

de même, on considère la fonction ϕ (fonction de mise à jour) tel que :
 $\phi : [0, 1] \times E \rightarrow [0, 1]$, définie par :

$$\phi(x, u) = \begin{cases} x & \text{si } u \in [0, P_{x_1}[\\ y & \text{si } u \in [P_{x_1}, P_{x_1} + P_{x_2}[\\ \vdots & \\ h & \text{si } u \in [\sum_{l=x}^{h-1} P_{x_l}, \sum_{l=x}^h P_{x_l}[\\ \vdots & \\ k & \text{si } u \in [\sum_{l=1}^{k-1} P_{x_l}, 1] \end{cases}$$

□

Proposition 5. *si $X_{n+1} = \phi(X_n, U_{n+1})$ alors on a :*

$$P[X_{n+1} = y / X_n = x] = P_{x,y}$$

Démonstration.

$$\begin{aligned}
P[X_{n+1} = h / X_n = x] &= P[\phi(X_n, U_{n+1}) = h / X_n = x] = P[\phi(X, U_{n+1}) = h] \\
&= \int_0^1 \mathbb{I}_{(\phi(x,u)=h)} d\mu = \mu(\phi(x, u) = h) \\
&= \sum_{l=x}^h P_{x_l} - \sum_{l=x}^{h-1} P_{x_l} = P_{x_h}
\end{aligned}$$

or la deuxième égalité est vraie car on a U_{n+1} est indépendante de (U_0, \dots, U_n)
et comme la suite (U_0, \dots, U_n) est indépendante de X_n (hypothèse), donc U_{n+1}
est indépendante de X_n □

Application Cette méthode donne un moyen de la simuler les trajectoires d'une chaîne de Markov.

A l'aide d'une fonction ψ , nous avons X_0 et à l'aide de ϕ nous avons X_n à partir de X_{n+1} , si on donne une suite de variables aléatoires $(U_n)_n$ iid de loi uniforme sur $[0, 1]$:

on a : $X_0 = \psi(U_0)$ on pose $X_n = \phi(X_{n+1}, U_n) \forall n \geq 0$

on affiche X_n .

L'algorithme donne alors une trajectoire (X_0, X_1, \dots, X_n) de $(X_n)_n$ a n pas

Exemple 1. Soit $E = \{x, y, z\}$ un espace des états fini $(X_n)_{n \geq 0}$ une chaîne sur E de loi initiale $\mu^0 = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ et de matrice de transition P tel que

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

Soit $(U_i)_{0 \leq i \leq 5}$ une suite de variables aléatoires iid de loi uniforme $U[0, 1]$

On a :

$$\psi(U) = \begin{cases} x & \text{si } u \in [0, \frac{1}{4}[\\ y & \text{si } u \in [\frac{1}{4}, \frac{3}{4}[\\ z & \text{si } u \in [\frac{3}{4}, 1] \end{cases}$$

D'autre part

$$\phi(x, u) = \begin{cases} x & \text{si } u \in [0, \frac{1}{2}[\\ y & \text{si } u = \frac{1}{2} \\ z & \text{si } u \in [\frac{1}{2}, 1] \end{cases}$$

et

$$\phi(y, u) = \begin{cases} x & \text{si } u \in [0, \frac{1}{3}[\\ y & \text{si } u \in [\frac{1}{3}, \frac{2}{3}[\\ z & \text{si } u \in [\frac{2}{3}, 1] \end{cases}$$

et nous avons $\forall u \in [0, 1]$, $\phi(z, u) = x$ d'où, on obtient l'échantillon suivant :

$$X_0 = \psi(U_0) ; X_1 = \phi(X_0, U_1)$$

$$X_2 = \phi(X_1, U_2) ; X_3 = \phi(X_2, U_3)$$

$$X_4 = \phi(X_3, U_4) ; X_5 = \phi(X_4, U_5).$$

Exemple 2. Simulation de temps de Gothenburg On considère la chaîne de Markov dans l'espace des états est $\{s_1, s_2\}$ où

$s_1 = \text{pluie}$ $s_2 = \text{ensoleillement}$ la matrice de transition est

$$\begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

Supposons que nous commençons la chaîne de Markov par un jour pluvieux de sorte que $\mu^{(0)} = (1, 0)$ Pour simuler cette chaîne ci-dessus en utilisant le schéma suivant la fonction initiale $\psi(U) = s_1 \forall x$ la fonction de mise à jour proposé par

$$\phi(s_1, u) = \begin{cases} s_1 & \text{si } u \in [0, \frac{3}{4}[\\ s_2 & \text{si } u \in [\frac{3}{4}, 1] \end{cases}$$

et

$$\phi(s_2, u) = \begin{cases} s_1 & \text{si } u \in [0, \frac{1}{4}[\\ s_2 & \text{si } u \in [\frac{1}{4}, 1] \end{cases}$$

d'où, on obtient l'échantillon suivant :

$$X_0 = \psi(U_0) , X_1 = \phi(X_0, U_1) , X_2 = \phi(X_1, U_2) , , X_3 = \phi(X_2, U_3)$$

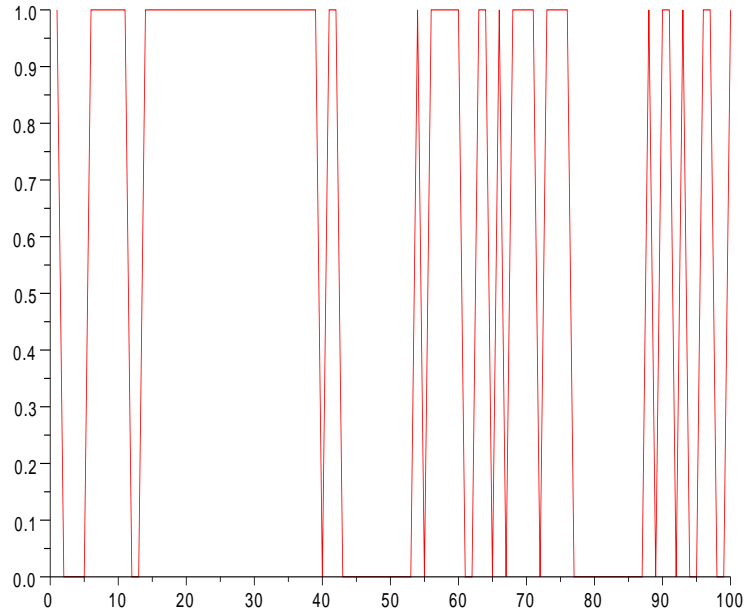


FIGURE 2.1 – Simulation des trajectoires de la chaîne de Markov modélisant le temps à Gothenburg sur 100 jours

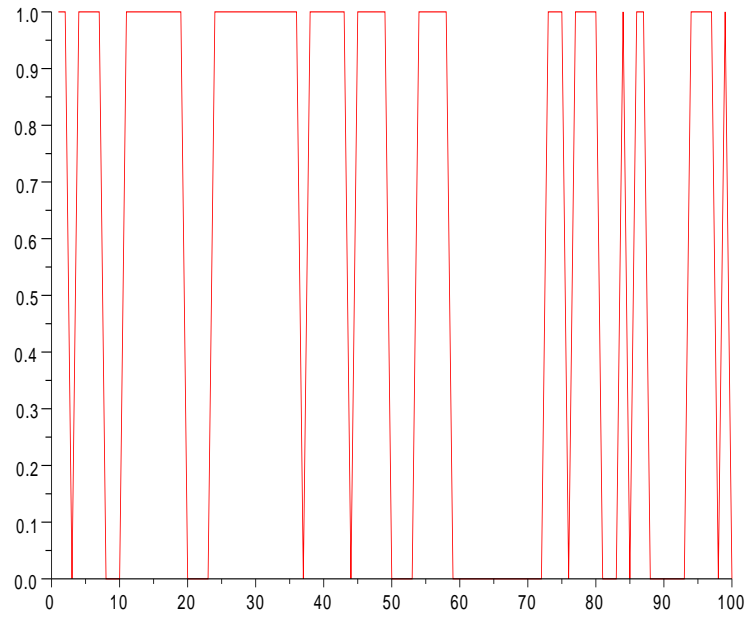


FIGURE 2.2 – Simulation des trajectoires de la chaîne de Markov modélisant le temps à Gothenburg sur 100 jours

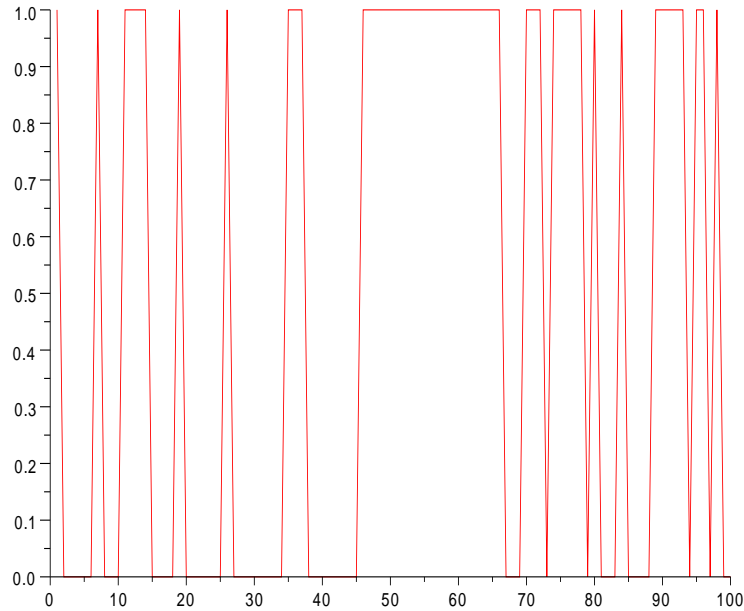


FIGURE 2.3 – Simulation des trajectoires de la chaîne de Markov modélisant le temps à Gothenburg sur 100 jours

Conclusion à ces observations On constate que souvent il pleut à Gothenburg mais des plages de beau temps sur une dizaine de jours apparaissent à chaque fois. Nous n'arrivons pas à expliquer ce phénomène. D'autant que la loi stationnaire de la chaîne montre que le beau temps et le temps pluvieux sont équiprobables.

2.3 Chaîne de Metropolis

Soit X une *v.a* de loi π . On veut un échantillon de cette loi (ou de X). On suppose qu'on dispose d'une chaîne de Markov $(X_n)_n$, de matrice de transition Ψ . L'algorithme de Metropolis est basé sur la construction d'une chaîne de Markov de loi stationnaire π . Nous décrivons cette méthode, ci-dessous.

Théorème 3. *Soit E un ensemble finie et π une loi de probabilité sur E , $(X_n)_n$ une chaîne de Markov sur E de probabilité de transition $(\psi(x, y))_{x, y \in E}$.
1) Si ψ est symétrique alors la chaîne de Markov sur E de matrice de transition $P = (P(x, y))_{x, y \in E}$ tel que*

$$P(x, y) = \begin{cases} \psi(x, y)[1 \wedge \frac{\pi_y}{\pi_x}] & \text{si } x \neq y \\ 1 - \sum_{z \neq x} \psi(x, z)[1 \wedge \frac{\pi_z}{\pi_x}] & \text{si } x = y \end{cases}$$

admet π comme loi stationnaire.

2) Si ψ est quelconque, la chaîne de Markov sur E de matrice de transition P tel que

$$P(x, y) = \begin{cases} \psi(x, y) \frac{[\frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} \wedge 1]}{[\frac{\pi_z \psi(z, x)}{\pi_x \psi(x, z)} \wedge 1]} & \text{si } x \neq y \\ 1 - \sum_{z \neq x} \psi(x, z) \frac{[\frac{\pi_z \psi(z, x)}{\pi_x \psi(x, z)} \wedge 1]}{[\frac{\pi_z \psi(z, x)}{\pi_x \psi(x, z)} \wedge 1]} & \text{si } x = y \end{cases}$$

admet π comme loi stationnaire.

Démonstration. Pour montrer que π est une loi stationnaire pour P on montre que π est réversible pour P .

π réversible si $\pi_x P(x, y) = \pi_y P(y, x)$ Si ψ symétrique. Pour $y \neq x$

$$\pi_x P(x, y) = \psi(x, y) [1 \wedge \frac{\pi_y}{\pi_x}] \dots (1)$$

$$\pi_y P(y, x) = \pi_y \psi(y, x) [1 \wedge \frac{\pi_x}{\pi_y}] \dots (2)$$

Si

$$\frac{\pi_y}{\pi_x} < 1$$

Nous avons

$$(1) = \pi_x \psi(x, y) \frac{\pi_y}{\pi_x} = \psi(x, y) \pi_y$$

et $(2) = \pi_y \psi(y, x) 1$ car

$$\frac{\pi_y}{\pi_x} < 1 \Rightarrow \frac{\pi_x}{\pi_y} > 1$$

.

$$\frac{\pi_y}{\pi_x} > 1$$

Nous avons

$$(1) = \pi_x \psi(x, y)$$

et

$$(2) = \pi_y \psi(y, x) \frac{\pi_x}{\pi_y} = \psi(y, x) \pi_x$$

Car

$$\frac{\pi_y}{\pi_x} > 1 \Rightarrow \frac{\pi_x}{\pi_y} < 1$$

donc $(1) = (2)$ car $\psi(x, y) = \psi(y, x)$ et donc π est réversible.

Cas où ψ est non symétrique :

π est réversible si $\pi_x P(x, y) = \pi_y P(y, x)$. Nous avons, pour $x \neq y$

$$\pi_x P(x, y) = \pi_x \psi(x, y) \left[\frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} \wedge 1 \right] \dots (*)$$

et

$$\pi_y P(y, x) = \pi_y \psi(y, x) \left[\frac{\pi_x \psi(x, y)}{\pi_y \psi(y, x)} \wedge 1 \right] \dots (**)$$

Si

$$\frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} < 1,$$

nous avons

$$(*) = \pi_x \psi(x, y) \frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} = \pi_y \psi(y, x)$$

et

$$(**) = \pi_y \psi(y, x) 1$$

car

$$\frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} < 1 \Rightarrow \frac{\pi_x \psi(x, y)}{\pi_y \psi(y, x)} > 1$$

Donc $(*) = (**)$ dans ce cas.

Si

$$\frac{\pi_y \psi(y, x)}{\pi_x \psi(x, y)} > 1,$$

un calcul analogue au précédent établit l'égalité recherchée et montre donc que π est réversible par rapport à la matrice P . \square

Conclusion Si π est une loi on sait associer à π une chaîne de Metropolis admetant π comme loi stationnaire. On peut approximer π par la loi de X_n . Soit \mathfrak{S} = ensemble des lois de probabilité dans E sur \mathfrak{S} . Soit d_{TV} la distance définie pour $\mu, \nu \in \mathfrak{S}$ par :

$$d_{TV}(\mu, \nu) = \sup_{A \in E} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$$

Nous avons alors pour toute loi initiale $\mu^{(0)}$, on a :

$$\mu^{(n)} \xrightarrow{d_{TV}} \pi \Leftrightarrow d_{TV}(\mu^{(n)}, \pi) \rightarrow 0 \quad n \rightarrow \infty$$

avec

$$\mu^{(n)} = \mathcal{L}(X_n)$$

Chapitre 3

Echantillonnage de Gibbs et Applications à la Physique

3.1 Introduction

Rappelons qu'un graphe $G = (V, E)$ est la donnée d'un ensemble V appelé l'ensemble des sommets et d'un ensemble E appelé ensemble des arêtes. Une arête est un chemin entre deux sommets adjacents.

Exemple 1. Soit G le graphe défini par :
 $V = \{v_1, v_2, v_3, v_4\}$ et $E = \{(v_1, v_2), (v_2, v_3), (v_1, v_3), (v_1, v_4), (v_3, v_4)\}$ On représente G par le schéma suivant :

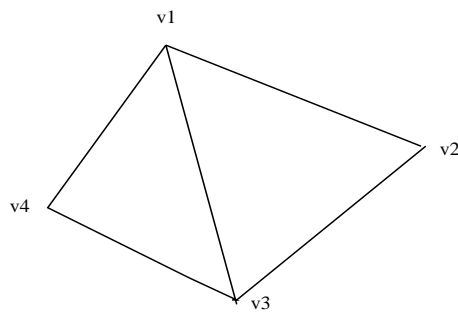


FIGURE 3.1 – exemple de graphe

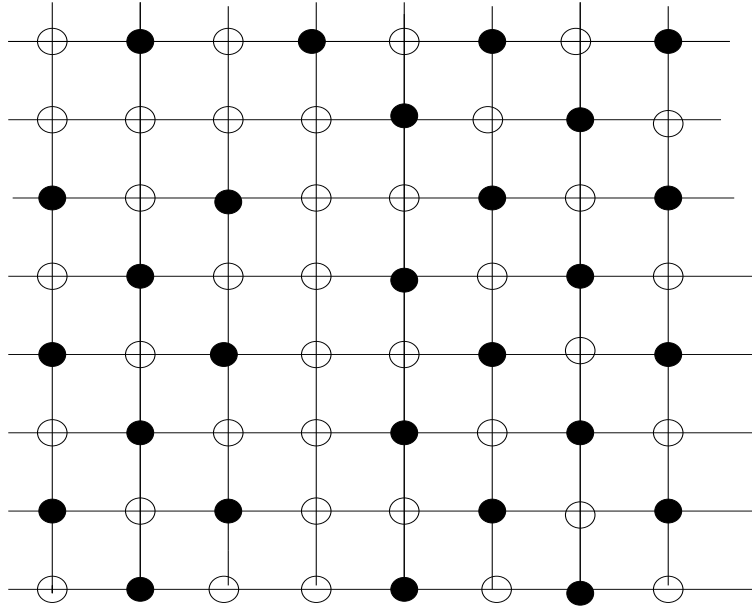


FIGURE 3.2 – exemple de configuration observable dans un graphe

Soit $G = (V, E)$ un graphe. On associe à chaque sommet v le nombre 0 ou 1, suivant que les sommets présentent oui ou non une propriété donnée (c'est le cas si les sommets sont occupés ou non par des particules d'un type donné). La règle est que deux sommets voisins (i.e. reliés par une arête) ne peuvent pas avoir la propriété ou être occupés tous les deux. On voudrait connaître le nombre moyen de sommets occupés par exemple. Pour cela nous avons besoin de quelques définitions.

Définition 1. Soit $G = (V, E)$ un graphe. L'application

$$x : V \rightarrow \{0, 1\}$$

est appelée une configuration. Soit

$$\{0, 1\}^V = \{x : V \rightarrow \{0, 1\}\}$$

Si $x \in \{0, 1\}^V$ est tel que $\forall v \in V$ si w est le voisinage de v noter par $w \sim v$ alors $x(v) = 1$ et $x(w) \neq 1$ donc x est une configuration qui est dite réalisable.

Soit Z_G le nombre de configurations réalisables associé au graphe G et

$$\mu_G(\xi) = \frac{1}{Z_G} \text{ pour } \xi \in \{0, 1\}^V, \xi \text{ observable}$$

$$\mu_G(\xi) = 0 \text{ si } \xi \text{ est non observable}$$

Problème Si on choisit une configuration réalisable on peut se poser la question savoir quel est le nombre moyen de 1 dans cette configuration ?.

Soit $n(\xi)$ = nombre 1 de la configuration ξ , et X une configuration choisie selon μ_G .

La valeur moyenne

$$\mathbb{E}[n(X)] = \sum_{\xi} n(\xi) \mu_G(\xi) = \frac{1}{Z_G} \sum_{\xi \text{ observable}} n(\xi)$$

Si le graphe est 8×8 , il ya $1,8 \times 10^{19}$ configurations ce qui fait l'impossibilité de ce calcul. On recourt alors aux simulations. Si on sait simuler une configuration aléatoire X suivant μ_G . On simule (X_1, \dots, X_n) , et on calcule $n(X_1), n(X_2), \dots, n(X_n)$ on sait d'après la loi de grand nombre que :

$$\frac{1}{n} \sum_{i=1}^n n(X_i) \rightarrow \mathbb{E}(n(X)) \text{ } x \rightarrow \infty, \text{ ps}$$

une estimation de $\mathbb{E}[n(X)]$ et donc

$$\frac{1}{n} \sum_{i=1}^n n(X_i)$$

le probleme est donc de simuler une configuration aléatoire X .

Proposition 1. Soit X une v.a à valeurs dans $S = \{s_1, \dots, s_n\}$ et de loi π tel que $\pi = (\pi_i)_{i=1}^n$. Si U est une v.a uniforme sur $[0, 1]$, soit $\psi : [0, 1] \rightarrow S$ la fonction de mise à jour définie par

$$\psi(x) = \begin{cases} s_1 & si & x \in [0, \pi_1[; \\ s_2 & si & x \in [\pi_1, \pi_1 + \pi_2[\\ \vdots & & \\ s_j & si & x \in [\sum_{i=1}^{j-1} \pi_i, \sum_{i=1}^j \pi_i] \\ \vdots & & \end{cases}$$

Alors la v.a $Y = \psi(U)$ est de même loi que X et donc $Y = X$ ps

Démonstration.

$$P[\psi(U) = s_j] = P\left[\sum_{i=1}^{j-1} \pi_i \leq U \leq \sum_{i=1}^j \pi_i\right] = \pi_j = P[X = s_j]$$

□

conclusion En pratique cette methode est difficilement utilisable sauf si l'espace S est petit. Pour le Hard-core model c'est pour construire ψ deja c'est trop long, car l'espace $S = \{0, 1\}^V$ est trop important.

Soit $G = \{V, E\}$ un graphe et $S = \{\xi \in \{0, 1\}^V, \xi \text{ est observable}\}$. notre but est de construire une chaîne de Markov $(X_n)_n$ d'espace des états S et admettant comme loi stationnaire la loi μ_G (loi uniforme) sur S . Ce problème est resolu par la proposition suivante.

Proposition 2. (*Algorithme de Gibbs*)

Soit $(X_n)_n$ la chaîne de Markov definie par l'algorithme suivant :

1. On choisit un sommet $v \in V$.
 2. On jette une piece reguliere.
 3. Si face sort, et tous les voisins w de v sont tel que $X_n(w) = 0$, alors on pose $X_{n+1}(v) = 1$, sinon $X_{n+1}(v) = 0$.
 4. Enfin $\forall w \neq v, X_{n+1}(w) = X_n(w)$.
- Alors $(X_n)_n$ est une chaîne de Markov apériodique, irreductible et admetant la loi stationnaire $\mu_Z = \frac{1}{Z_G}$ si Z_G le nombre total de configuration réalisable.

Démonstration. Pour toute configuration réalisable ξ

$$\xi \rightarrow \xi \text{ donc, } P_{\xi\xi} > 0$$

donc $(X_n)_n$ aperiodique.

$(X_n)_n$ est irreductible car chaque transition ne touche qu'un sommet. Donc si ξ et ξ' sont deux configuration réalisables, on cherche tous les sommets v ou $\xi(v) \neq \xi'(v)$ et on change tous les sommets un à un par la chaîne $(X_n)_n$. La chaîne va donc passer de ξ à ξ' . Cela montre que tous les états comunique.

Enfin μ_G est une loi stationnaire car μ_G est réversible par rapport à $(X_n)_n$.

En effet, soit ξ et ξ' deux configurations observables et $d = d(\xi, \xi')$ le nombre de sommets v où $\xi(v) \neq \xi'(v)$

Cas $d = 0$

$P_{\xi, \xi'} = P_{\xi', \xi}$, $\mu_G(\xi) = \mu_G(\xi')$ Donc immédiatement nous avons

$$\mu_G(\xi)P_{\xi, \xi'} = \mu_G(\xi')P_{\xi', \xi} \dots (1)$$

Cas $d \geq 2$:

On ne peut pas aller en un pas de ξ à ξ' et inversement et donc $P_{\xi, \xi'} = P_{\xi', \xi} = 0$. Donc (1) est valable.

Cas $d = 1$:

Soit v tel que $\xi(v) \neq \xi'(v)$

Pour passer de ξ à ξ' (ou de ξ' à ξ). On choisit un sommet dans V suivant la loi uniforme, avec $|V| = n$ et on tire un pile ou face sont exactement tous les sommets voisins de v sont tel que $\xi(v) = 0$. Le choix de v se fait avec la probabilité

$$\frac{1}{n}$$

et Face a la probabilité

$$\frac{1}{2}$$

Donc la transition de ξ à ξ' se fait avec la la probabilité

$$\frac{1}{2n}.$$

Par suite (1) devient

$$\frac{1}{Z_G} \frac{1}{2n} = \frac{1}{Z_G} \frac{1}{2n}$$

car $P_{\xi, \xi'} = P_{\xi', \xi}$. Donc (1) est vérifiée et $\mu_G = \frac{1}{Z_G}$ est bien réversible et donc stationnaire pour $(X_n)_n$.

□

3.2 Introduction au modèle d'Ising

Soit $G = (V, E)$ un graphe et $S = \{-1, 1\}^V$. Si $\sigma \in S$, σ est une configuration. Pour $v \in V$, si $\sigma \in S$, $\sigma(v) = \pm 1$. On interprète la valeur $\sigma(v) = 1$

lorsque le spin est orienté vers le haut et $\sigma(v) = -1$ lorsque le spin est orienté vers le bas.

On appelle modèle d'Ising la repartition des spins sur le graph G . On se limite au cas dit du "modèle du voisin le plus proche" (nearest neighboring model). Pour $\sigma \in S$, l'énergie interne associée à σ , H_σ est définie par

$$H_\sigma = - \sum_{v,w \in V \text{ et } v \sim w} \sigma(v)\sigma(w)$$

Dans ce modèle la probabilité μ_G , définie sur S par

$$\mu_G(\sigma) = \frac{1}{Z_\beta} e^{-\beta H_\sigma}$$

où Z_β est une constante de normalisation et $\beta = \frac{1}{kT}$ ou T est la température et k = constante de Boltzmann, est appelée la loi de Gibbs. Z_G est définie par $\sum_{\sigma \in S} \mu_G(\sigma) = 1$ et donc

$$Z_\beta = \sum_{\sigma \in S} e^{-\beta H_\sigma}.$$

L'Algorithme de Gibbs pour le modèle d'Ising est donné par la proposition suivante :

Proposition 3. Soit $G = (V, E)$ un graphe et $S = \{-1, 1\}^V$ (Cf. [2]) La chaîne de Markov sur S de probabilités de transitions, pour σ et $\sigma' \in S$.

$$P(\sigma, \sigma') = \frac{1}{n} \sum_{v \in V} \frac{e^{\beta \sigma'(v) S(\sigma, v)}}{e^{\beta \sigma'(v) S(\sigma, v)} + e^{-\beta \sigma'(v) S(\sigma, v)}} \times 1_{\{\sigma(w) = \sigma'(w) \text{ pour } w \neq v\}}$$

admet la distribution de Gibbs, μ_G comme loi stationnaire.

Conclusion Comme nous disposons d'un algorithme de simulation d'une chaîne de Markov de loi stationnaire π , la proposition fournit donc la moyenne de construire un échantillon (X_1, \dots, X_n) de configurations de loi μ_G . Ce qui permet de fournir les statistiques d'autant pour ce modèle

Conclusion

Dans ce mémoire nous avons présenté quelques algorithmes de simulation. Les plus importants sont ceux de Metropolis et Gibbs. Il manque les programmes pour mettre en oeuvre ces algorithmes. Une question importante est le temps d'exécution de ces programmes. En particulier il s'agit de savoir que ce sont des algorithmes polynomiaux. cela est d'autant plus pertinent pour l'échantillonnage de Gibbs dans le cadre du modèle des sphères lourdes ou dans le cadre du modèle d'Ising. Cela a retenu notre attention compte tenu de la demande en thermodynamique et en physique.

C'est un domaine de travail qui s'ouvre, qu'il faut bien aborder et approfondir. Les seules réponses que nous pouvons donner pour l'instant sont celles d'expliquer le pourquoi des algorithmes proposés (Metropolis, Gibbs,...).

Programmes utilisés pour les simulations

Nous donnons ci-après les programmes utilisés pour la simulation de la Marche aléatoire sur \mathbb{Z} et la chaîne de Markov modélisant le temps qu'il fait à Gothenburg.

Programme Scilab de simulation de la Marche aléatoire sur \mathbb{Z}

```
val=0.5 ; n=500 ;  
y(1)=0 ;  
for i=2 :n, u=rand() ;  
if u<val, x(i)=1 ; else x(i)=-1 ;  
end  
y(i)=y(i-1)+x(i) ;  
end  
j=[1 :1 :n], y=y(j) ; plot(j,y)
```

Programme Scilab de simulation du temps à Gothenburg

```
n=100 ;  
temps=zeros(1,n) ;  
temps(1)= 1 ;  
for i=1 :n-1  
u=rand() ;  
if temps(i)== 1  
if u<=0.75  
temps(i+1)= 1 ;  
else  
temps(i+1)= 0 ;  
end  
else
```

```
if u<=0.25,  
temps(i+1)= 1 ;  
else  
temps(i+1)= 0 ;  
end  
end  
end j=(1 :n)'; z=temps(j); plot2d(j,z,style=5)
```

Bibliographie

- [1] Samuel Karlin and Howard M. Taylor *A first course in stochastic processes*
Academic press, INC, 1975

- [2] Ycart B. *Algorithmes markoviens*
Département de l'ingénieur de mathématique. Université de Chile. Decembre 1997.

- [3] Drif Zina Mémoire de Master *Algorithme de simulation de PROPP-WILSON*
Département de Maths, UMMTO, 2013.

- [4] Arezki Ourdia Mémoire de Master *Fractions continus et systèmes dynamiques*
Département de Maths. UMMTO, 2013.

- [5] Levin David A. and al. *Markov chains and Mixing Times*
Univesity of Oregon.

- [6] Hägström M,O. *Finite Markov chaîne and algorithmic application*
Cambridge University Press, 2002.

- [7] Benaïm M. et El Karoui N. *Premenade aléatoire, chaîne de Markov et simulation; martingales et stratégies*
Edition de l'école polytechnique, Novembre 2007.

- [8] Muller Scott L. and Childers Donald G. *Probability and Random processes with application to signal processing and communication*

Elsevier Academic Press, 2004.

- [9] Diaconis, P and Saloff-coste *What do we know about the Metropolis algorithm ?*
Journal of computer and system sciences 57,20-36(1998)