

Cours de Probabilités

(M1 de Mathématiques. U.d.S.)

J. FRANCHI

Deuxième semestre 2010-2011. 52 heures de cours intégrés.

Contenu

I. Vecteurs Gaussiens	page 2
1) Vecteurs aléatoires	page 2
2) Variables réelles gaussiennes (rappels)	page 2
3) Variables gaussiennes vectorielles	page 3
II. Marches aléatoires sur \mathbb{Z}^d	page 6
1) Marches aléatoires et temps d'arrêt	page 6
2) Temps de retour en 0	page 7
3) Récurrence des marches aléatoires simples	page 9
4) Récurrence des autres marches aléatoires	page 11
5) Retour en 0 et loi de l'Arcsinus pour la marche simple sur \mathbb{Z}	page 12
III. Chaînes de Markov	page 16
1) La propriété de Markov	page 16
2) Classification des états	page 21
3) Mesures invariantes des chaînes finies	page 24
4) Probabilité invariante	page 26
5) Ergodicité	page 29
IV. Martingales à temps discret	page 31
1) Surmartingales, sousmartingales, théorème d'arrêt	page 31
2) Inégalités	page 33
3) Convergence	page 35
V. Processus de Poisson	page 38
1) Processus de Poisson homogènes en dimension 1	page 39
2) Autres processus de Poisson	page 44
VI. Files d'attente	page 45
1) Files à un serveur	page 45
2) Files $\mathcal{G}_p/\mathcal{G}_q/1$	page 47
3) Files $\mathcal{G}_p/F_S/1$	page 49

1 Vecteurs Gaussiens

1.1 Vecteurs aléatoires

On appelle vecteur aléatoire toute v.a. à valeurs dans \mathbb{R}^d .

On dit qu'il est dans L^p lorsque sa norme l'est (pour tout $p \in [1, \infty]$).

Exercice n° 1.1.1 a) Vérifier que la définition ci-dessus ne dépend pas de la norme choisie.

b) Montrer qu'un vecteur aléatoire est dans L^p ssi ses coordonnées (dans n'importe quelle base) sont dans L^p .

On étend par linéarité la définition de l'espérance pour les variables réelles intégrables aux vecteurs aléatoires intégrables. Ainsi, dans n'importe quelle base, l'espérance d'un vecteur aléatoire intégrable V est le vecteur ayant pour coordonnées les espérances des coordonnées de V .

Définition 1 On appelle matrice de covariance (ou de dispersion) d'un vecteur aléatoire V de carré intégrable (i.e. dans L^2) la matrice $K_V := \mathbb{E}\left((V - \mathbb{E}(V)) \times^t (V - \mathbb{E}(V))\right)$, où les vecteurs sont considérés comme des colonnes et tV désigne le transposé de V .

Exercice n° 1.1.2 a) Vérifier que $K_V = \mathbb{E}(V \times^t V) - \mathbb{E}(V) \times^t \mathbb{E}(V) = ((Cov(V_i, V_j)))_{i,j}$.

b) Vérifier que pour tout $u \in \mathbb{R}^d$ on a $Var({}^tuV) = {}^tuK_Vu = \sum_{1 \leq i, j \leq d} u_i u_j Cov(V_i, V_j)$.

c) Montrer que K_V est une matrice symétrique positive.

d) Montrer que K_V est inversible ssi il n'existe pas d'hyperplan de \mathbb{R}^d contenant p.s. V .

e) Soit M une matrice de $\mathcal{M}_{n,d}(\mathbb{R})$. Calculer $\mathbb{E}(MV)$ et K_{MV} . Même question pour AV , si A est une application affine de \mathbb{R}^d dans \mathbb{R}^n .

1.2 Variables réelles gaussiennes (rappels)

Définition 2 Une variable aléatoire réelle est dite gaussienne centrée réduite ou normale centrée réduite ou gaussienne standard lorsqu'elle admet la densité :

$t \mapsto \exp(-t^2/2)/\sqrt{2\pi}$. Une variable aléatoire réelle X est dite gaussienne ou normale lorsqu'il existe $m \in \mathbb{R}$ et $\sigma > 0$ tels que $(X - m)/\sigma$ soit normale centrée réduite. On dit alors que la loi de X est $\mathcal{N}(m, \sigma^2)$.

Exercice n° 1.2.1 a) Vérifier que $m = \mathbb{E}(X)$ et $\sigma^2 = Var(X)$, et que $X \in \cap_{p < \infty} L^p$.

b) Vérifier que $\frac{A^2 e^{-x^2/2}}{(A^2+1)x} \leq \int_x^\infty e^{-t^2/2} dt \leq \frac{e^{-x^2/2}}{x}$ sur $[A, \infty[$, pour tout $A > 0$. Donner un équivalent de $\mathbb{P}(X > x)$ en $+\infty$.

c) Montrer que la densité de X est $t \mapsto \exp(-(t - m)^2/2\sigma^2)/\sigma\sqrt{2\pi}$.

d) Montrer que la transformée de Fourier de X est donnée par $\mathbb{E}(e^{itX}) = e^{itm - \sigma^2 t^2/2}$, et

que cela caractérise la loi de X .

e) Montrer que toute combinaison linéaire (non triviale) de v.a. normales indépendantes est normale. Donner un contreexemple simple s'il n'y a pas indépendance.

Exercice n° 1.2.2 Notons $\{X_n | n \in \mathbb{N}^*\}$ une suite de v.a.r. indépendantes gaussiennes standard. Etudier les convergences en loi et en probabilité de $\frac{nX_1 + (n-1)X_2 + \dots + X_n}{n\sqrt{n}}$.

1.3 Variables gaussiennes vectorielles

Définition 3 Une variable aléatoire V à valeurs dans \mathbb{R}^d est dite *gaussienne* ou *normale* lorsque pour tout $u \in \mathbb{R}^d$ ${}^t u V$ est une v.a.r. normale ou p.s. constante.

On note $\mathcal{N}(m, K)$ la loi d'un vecteur gaussien d'espérance m et de matrice de covariance K . Une probabilité (resp. une densité) est dite *gaussienne* lorsqu'elle est la loi (resp. la densité) d'un vecteur gaussien.

Exercice n° 1.3.1 a) Vérifier que si V est un vecteur gaussien de \mathbb{R}^d et si A est une application affine de \mathbb{R}^d dans \mathbb{R}^n , alors AV est un vecteur gaussien.

b) Montrer que si V est un vecteur gaussien, alors ses coordonnées dans n'importe quelle base sont gaussiennes (ou p.s. constantes).

c) Soient X gaussienne standard, et ε indépendante de X et uniforme sur ± 1 . Montrer que εX est gaussienne standard, mais que le vecteur $(X, \varepsilon X)$ n'est pas gaussien.

d) Montrer que si le vecteur aléatoire V a ses coordonnées gaussiennes et indépendantes, alors il est gaussien.

e) Montrer que la transformée de Fourier d'un vecteur gaussien V est donnée par

$$\mathbb{E}(e^{i {}^t u V}) = \exp\left(i {}^t u \mathbb{E}(V) - \frac{1}{2} {}^t u K_V u\right), \text{ pour tout } u \in \mathbb{R}^d, \text{ et qu'elle caractérise sa loi.}$$

f) Montrer que les coordonnées (dans la base canonique de \mathbb{R}^d) d'un vecteur gaussien sont indépendantes ssi la matrice de covariance est diagonale, et donc ssi elles sont non corrélées. Vérifier que c'est faux si le vecteur n'est pas gaussien (même si ses coordonnées sont gaussiennes).

Lemme 1 Soit K une matrice réelle symétrique positive, de format $d \times d$ et de rang $r \in \{0, \dots, d\}$. Alors

1) Il existe une matrice réelle symétrique positive S de format $d \times d$ telle que $K = S^2$.

2) Il existe une matrice réelle M de format $d \times r$ et de rang r telle que $K = M {}^t M$.

Preuve Il existe une matrice orthogonale P et une matrice diagonale positive D telles que $K = P D P^{-1}$. $S := P \sqrt{D} P^{-1}$ fournit une solution pour 1). Permutant les vecteurs propres de K , ce qui revient à changer de matrice orthogonale P , on se ramène au cas où les r premiers coefficients diagonaux de D sont > 0 et les autres nuls. Alors si Δ est égale à la matrice formée des r premières colonnes de \sqrt{D} , on voit aussitôt que $M := P \Delta$ est solution de 2) : $K_{ij} = \sum_{k=0}^d p_{ik} d_k p_{jk} = \sum_{k=0}^r p_{ik} \sqrt{d_k} p_{jk} \sqrt{d_k} = (M {}^t M)_{ij}$. \diamond

Ce lemme permet d'exhiber un vecteur gaussien ayant une espérance et une matrice de covariance données à l'avance. Le résultat précis est le suivant.

Proposition 1 Soit K une matrice réelle symétrique positive, de format $d \times d$ et de rang r , et soit M une matrice réelle de format $d \times r$ telle que $K = M \times^t M$. Soit $m \in \mathbb{R}^d$. Soit V un vecteur de \mathbb{R}^r formé de coordonnées gaussiennes standard indépendantes. Alors $m + MV$ est un vecteur gaussien de loi $\mathcal{N}(m, K)$.

Corollaire 1 Si la matrice de covariance d'un vecteur gaussien d -dimensionnel de loi $\mathcal{N}(m, K)$ est inversible, alors ce vecteur admet la densité sur \mathbb{R}^d :

$$v \longmapsto (2\pi)^{-d/2} (\det K)^{-1/2} \exp\left(-\frac{1}{2} {}^t(v-m)K^{-1}(v-m)\right).$$

Preuve Soit V un vecteur de \mathbb{R}^d formé de coordonnées gaussiennes standard indépendantes. Soit M une matrice réelle de format $d \times d$ telle que $K = M \times^t M$. Alors M est inversible, et nous avons pour toute fonction-test f , par changement de variable :

$$\begin{aligned} \mathbb{E}(f(m+MV)) &= \int_{\mathbb{R}^d} f(m+Mx) \prod_{j=1}^d (2\pi)^{-1/2} e^{-x_j^2/2} dx = \int_{\mathbb{R}^d} f(m+Mx) (2\pi)^{-d/2} e^{-{}^t x x/2} dx \\ &= \int_{\mathbb{R}^d} f(v) (2\pi)^{-d/2} e^{-{}^t(M^{-1}(v-m)) \times (M^{-1}(v-m))/2} |\det(M^{-1})| dv \\ &= \int_{\mathbb{R}^d} f(v) (2\pi)^{-d/2} (\det K)^{-1/2} \exp\left(-\frac{1}{2} {}^t(v-m)K^{-1}(v-m)\right) dv. \quad \diamond \end{aligned}$$

Exercice n° 1.3.2 Soit V un vecteur gaussien d -dimensionnel, et pour $1 \leq j \leq k$ soient M_j une matrice réelle de format $d_j \times d$ et $V_j := M_j V$. Montrer que V_i et V_j sont indépendants ssi $M_i K_V {}^t M_j = 0$. Généraliser à l'indépendance de V_1, \dots, V_k .

Exercice n° 1.3.3 Soit $\{X_j | j \in \mathbb{N}^*\}$ une suite de v.a.i.i.d. $\mathcal{N}(0, 1)$. Soit $\{a_j | j \in \mathbb{N}\}$ une suite de réels tels que $a_j a_{j+1} = 0$ pour tout j et tels que $\sum_j a_j^2 < \infty$. Soit $Y_n := \sum_{j=1}^n a_{n-j} X_j$, pour tout $n \in \mathbb{N}^*$.

a) Etudier la convergence en loi de la suite Y_n . b) Les v.a. Y_n et Y_{n+1} sont-elles indépendantes ? c) Etudier la convergence dans L^2 de la suite Y_n .

Exercice n° 1.3.4 Soit $V := (V_0, \dots, V_d)$ un vecteur gaussien $(d+1)$ -dimensionnel, dont les coordonnées sont $\mathcal{N}(0, 1)$ et vérifient : $Cov(V_0, V_j) = p$ pour $1 \leq j \leq d$ et $Cov(V_i, V_j) = p^2$ pour $1 \leq i \neq j \leq d$, p étant un paramètre. Posons $W_j := (1-p^2)^{-1/2}(V_j - pV_0)$ pour $1 \leq j \leq d$. Déterminer successivement les lois de : (V_0, W_1, \dots, W_d) ; $S := \sum_{j=1}^d V_j$; S/V_0 .

Exercice n° 1.3.5 Soit $\{X_j | 1 \leq j \leq n\}$ une suite de v.a.r.i.i.d. $\mathcal{N}(m, \sigma^2)$. Soient $\bar{X} := (X_1 + \dots + X_n)/n$, $X'_j := X_j - \bar{X}$, et $S := ((X'_1)^2 + \dots + (X'_n)^2)/\sigma^2$. Préciser la loi de \bar{X} , de $(\bar{X}, X'_1, \dots, X'_n)$, montrer que \bar{X} et S sont indépendantes. Montrer que la loi de S est un χ^2 . (Utiliser une b.o.n. B de l'hyperplan $\sum_j x'_j = 0$, complétée en une b.o.n. de \mathbb{R}^n .)

Voici la version vectorielle du théorème limite central.

Théorème 1 Soit $\{X_j \mid j \in \mathbb{N}^*\}$ une suite de v.a.i.i.d. de carré intégrable à valeurs dans \mathbb{R}^d . Alors lorsque $n \rightarrow \infty$ la loi de $(X_1 + \dots + X_n - n\mathbb{E}(X_1))/\sqrt{n}$ converge vers $\mathcal{N}(0, K_{X_1})$.

Preuve L'énoncé étant clairement invariant par une translation sur la suite X , il suffit de considérer le cas où X_1 est centrée. Notons $\phi(v) := \mathbb{E}(e^{i v \cdot X_1})$ la transformée de Fourier de X_1 , qui est de classe C^2 , puisque X_1 est supposée de carré intégrable. La formule de Taylor-Young à l'origine s'écrit alors : $\phi(v) = 1 - \frac{1}{2} {}^t v K(X_1) v + o(\|v\|^2)$. Par indépendance, la transformée de Fourier de $\frac{X_1 + \dots + X_n}{\sqrt{n}}$ est égale à $\phi(v/\sqrt{n})^n$, qui vaut $\left[1 - \frac{{}^t v K(X_1) v + n o(\|v\|^2/n)}{2n}\right]^n$, et donc qui converge vers $\exp\left[-\frac{1}{2} {}^t v K(X_1) v\right]$, du fait de la convergence de $(1 + \frac{z}{n})^n$ vers e^z uniformément sur les compacts de \mathbb{C} .

En effet, si $|z| \leq A$, nous avons :

$$\begin{aligned} \left| e^z - \left(1 + \frac{z}{n}\right)^n \right| &\leq \sum_{k=0}^n \left(\frac{1}{k!} - \frac{C_n^k}{n^k}\right) |z|^k + \sum_{k=n+1}^{\infty} \frac{|z|^k}{k!} \leq \sum_{k=0}^n \left(\frac{1}{k!} - \frac{C_n^k}{n^k}\right) A^k + \sum_{k=n+1}^{\infty} \frac{A^k}{k!} \\ &= e^A - \left(1 + \frac{A}{n}\right)^n = e^A - e^{n \log(1 + \frac{A}{n})} \longrightarrow 0. \end{aligned}$$

Finalement, nous venons de montrer la convergence simple de la transformée de Fourier de $\frac{X_1 + \dots + X_n}{\sqrt{n}}$ vers celle de $\mathcal{N}(0, K_{X_1})$, ce qui suffit. \diamond

Exercice n° 1.3.6 Soit $(X_1, \dots, X_d, Y) = ({}^t X, Y)$ un vecteur gaussien, de matrice de covariance inversible $K := \begin{pmatrix} K_X & V \\ {}^t V & \sigma^2 \end{pmatrix}$, σ étant l'écart-type de Y . Notons (e_1, \dots, e_{d+1}) la base canonique de \mathbb{R}^{d+1} , et pour $u, v \in \mathbb{R}^{d+1}$: $\langle u, v \rangle := {}^t u K v$. Notons encore ${}^t a := (a_1, \dots, a_d)$ les coordonnées dans (e_1, \dots, e_d) de la projection orthogonale de e_{d+1} sur \mathbb{R}^d , relativement au produit scalaire $\langle \cdot, \cdot \rangle$.

Posons $X'_j := X_j - \mathbb{E}(X_j)$, pour $1 \leq j \leq d$, $X' := ({}^t X'_1, \dots, X'_d)$, et $Y' := Y - \mathbb{E}(Y) - {}^t a X'$.

- Vérifier que K_X est nécessairement inversible. Calculer a et $\text{Var}(Y')$ en fonction de K .
- Calculer la covariance de X'_j et Y' , et déduire en fonction de a l'espérance conditionnelle de Y' , puis celle de Y , par rapport à X .
- Vérifier, en calculant $\mathbb{E}(f \circ Y \mid X)$, que la loi conditionnelle de Y sachant X est gaussienne ; préciser ses paramètres.

Exercice n° 1.3.7 Soit $(X_1, X_2, X_3) \in \mathbb{R}^3$ un vecteur gaussien centré.

- Montrer que X_1 et X_2 sont indépendants ssi $\mathbb{E}[\exp(X_1 + X_2)] = \mathbb{E}[\exp(X_1)]\mathbb{E}[\exp(X_2)]$.
- Si $\mathbb{E}[\exp(X_1 + X_2 + X_3)] = \mathbb{E}[\exp(X_1)]\mathbb{E}[\exp(X_2)]\mathbb{E}[\exp(X_3)]$, peut-on conclure à l'indépendance de X_1, X_2, X_3 ? Justifier.

Exercice n° 1.3.8 Soient X, Y deux v.a.r.i.i.d. centrées réduites.

- Lorsqu'elles sont gaussiennes, pour quels réels a, b, c, d a-t-on $aX + bY$ et $cX + dY$ indépendantes ?

Supposons dorénavant $X - Y$ et $X + Y$ indépendantes ; on veut établir que X, Y sont nécessairement gaussiennes.

- b) Montrer que $\hat{X}(2t) = \hat{X}(t)^3 \hat{X}(-t)$ pour tout $t \in \mathbb{R}$.
 c) Montrer que \hat{X} ne s'annule pas.
 d) Dédurre que \hat{X} est réelle paire, et conclure. (Considérer une détermination continue de l'argument, puis $\varphi(t) := (\arg \hat{X}(t))/t$, et utiliser que X est réelle.)

2 Marches Aléatoires sur \mathbb{Z}^d

2.1 Marches aléatoires et temps d'arrêt

Définition 4 Soit $\{X_n | n \in \mathbb{N}^*\}$ une suite de v.a. à valeurs dans \mathbb{Z}^d , indépendantes, et de même loi μ . On appelle marche aléatoire (sur \mathbb{Z}^d) le processus $\{S_n | n \in \mathbb{N}\}$ à valeurs dans \mathbb{Z}^d défini par : $S_0 := 0$, et pour tout $n \in \mathbb{N}^*$: $S_n := X_1 + \dots + X_n$. La marche est dite simple lorsque $\mu(\{e_j\}) = \mu(\{-e_j\}) = 1/(2d)$, où (e_1, \dots, e_d) désigne la base canonique de \mathbb{Z}^d .

Définition 5 Soit pour tout $n \in \mathbb{N}^*$: $\mathcal{F}_n := \sigma(X_1, \dots, X_n) = \sigma(S_1, \dots, S_n)$ la tribu représentant l'information connue au temps n . On note aussi \mathcal{F}_0 la tribu triviale, et \mathcal{F}_∞ la tribu engendrée par $\cup_{\{n \in \mathbb{N}\}} \mathcal{F}_n$. On appelle temps d'arrêt toute v.a. N à valeurs dans $\overline{\mathbb{N}}$ telle que $\{N \leq n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$ (ou $\overline{\mathbb{N}}$). On pose alors : $\mathcal{F}_N := \{A \in \mathcal{F}_\infty | (\forall n \in \mathbb{N}) A \cap \{N \leq n\} \in \mathcal{F}_n\}$.

L'exemple fondamental de temps d'arrêt est le temps d'atteinte par le processus d'une certaine partie E de \mathbb{Z}^d : $\min\{n \in \mathbb{N} | S_n \in E\}$ est bien un temps d'arrêt. Un autre exemple, trivial, est fourni par les temps constants.

Exercice n° 2.1.0 Montrer que N est un temps d'arrêt ssi $\{N = n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$, et que $\mathcal{F}_N = \{A \in \mathcal{F}_\infty | (\forall n \in \mathbb{N}) A \cap \{N = n\} \in \mathcal{F}_n\}$.

Exercice n° 2.1.1 Soient N et N' deux temps d'arrêt tels que $N \leq N'$.

- a) Vérifier que \mathcal{F}_N est une tribu. b) Montrer que $\mathcal{F}_N \subset \mathcal{F}_{N'}$.
 c) Soit $A \in \mathcal{F}_N$. Montrer que $N1_A + N'1_{A^c}$ est un temps d'arrêt.
 d) Montrer que si Z est un processus adapté (i.e. : $Z_n \in \mathcal{F}_n$ pour tout n), alors $Z_N \in \mathcal{F}_N$.

Exercice n° 2.1.2 Soient N et N' deux temps d'arrêt.

- a) Vérifier que $\min\{N, N'\}$ et $\max\{N, N'\}$ sont aussi deux temps d'arrêt ; et $N + N'$?
 b) Montrer que les événements $\{N < N'\}$, $\{N \leq N'\}$, $\{N = N'\}$ appartiennent à $\mathcal{F}_{\min\{N, N'\}}$, puis que $\mathcal{F}_{\min\{N, N'\}} = \mathcal{F}_N \cap \mathcal{F}_{N'}$.

Exercice n° 2.1.3 Soit N un temps d'arrêt. Supposons que $\mathbb{E}(\|X_1\|)$ et $\mathbb{E}(N)$ sont finies. Montrer que $\mathbb{E}(S_N) = \mathbb{E}(X_1) \times \mathbb{E}(N)$. (Écrire $S_N = \sum_j X_j 1_{\{N < j\}^c}$).

Lemme 2 Soit N un temps d'arrêt. Conditionnellement à $\{N < \infty\}$, la suite $\{X_{N+n} | n \in \mathbb{N}^*\}$ est indépendante de \mathcal{F}_N , et a la même loi que $\{X_n | n \in \mathbb{N}^*\}$.

Preuve Fixons $A \in \mathcal{F}_N$, $k \in \mathbb{N}^*$, et $v_1, \dots, v_k \in \mathbb{Z}^d$. Nous avons

$$\begin{aligned} \mathbb{P}(A, N < \infty, X_{N+1} = v_1, \dots, X_{N+k} = v_k) &= \sum_{n \in \mathbb{N}} \mathbb{P}(A, N = n, X_{n+1} = v_1, \dots, X_{n+k} = v_k) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(A, N = n) \mathbb{P}(X_{n+1} = v_1, \dots, X_{n+k} = v_k) = \mathbb{P}(A, N < \infty) \prod_{j=1}^k \mu(v_j). \quad \diamond \end{aligned}$$

Quitte à prendre l'image de Ω par X , on peut toujours se placer sur l'espace canonique, i.e. considérer que $\Omega = (\mathbb{Z}^d)^{\mathbb{N}^*}$, $\mathcal{F} = \mathcal{F}_\infty$, de sorte que X devient la suite des coordonnées : $X_n(\omega) = \omega_n$, $\forall n \in \mathbb{N}$. \mathbb{P} demeure la loi de la marche S , déduite de $(X, \mu^{\otimes \mathbb{N}^*})$.

Nous avons alors la notion de décalage, ou shift : c'est l'application θ de Ω dans Ω définie par $(\theta\omega)_n := \omega_{n+1}$, pour tout $n \in \mathbb{N}^*$, id est $X_n \circ \theta = X_{n+1}$ (et $S_n \circ \theta = S_{n+1} - S_1$). Puis $\theta^0 := Id$, $\theta^1 := \theta$, $\theta^{k+1} := \theta \circ \theta^k$ définit θ^k pour tout $k \in \mathbb{N}$, qui vérifie : $(\theta^k\omega)_n = \omega_{n+k}$, pour tout $n \in \mathbb{N}^*$, id est $X_n \circ \theta^k = X_{n+k}$ (et $S_n \circ \theta^k = S_{n+k} - S_k$). Enfin si N est un temps d'arrêt, on pose : $\theta^N := \theta^n$ sur $\{N = n\}$, et $\theta^N\omega := \delta$ si $N(\omega) = \infty$, où $\delta \notin \Omega$ est un point qu'on rajoute à Ω , et qu'on nomme cimetière.

Le lemme 2 ci-dessus se réécrit : *Conditionnellement à $\{N < \infty\}$,*

la suite $\{S_n \circ \theta^N \mid n \in \mathbb{N}\}$ est indépendante de \mathcal{F}_N , et a la même loi que $\{S_n \mid n \in \mathbb{N}\}$.

Nota Bene : $S_n \circ \theta^N = S_{N+n} - S_N$.

Exercice n° 2.1.4 Soient N et N' deux temps d'arrêt. Montrer que $N + N' \circ \theta^N$ est un temps d'arrêt. Interpréter, lorsque N et N' sont des temps d'atteinte. Évaluer $\theta^{N'} \circ \theta^N$.

Exercice n° 2.1.5 Supposons $\mathbb{E}(X_1^+) < \infty$, $\mathbb{E}(X_1) < 0$; notons $M := \sup\{S_j \mid j \in \mathbb{N}^*\}$ puis $M_n := \max\{S_j \mid 1 \leq j \leq n\}$ pour $n \in \mathbb{N}^*$. On veut établir que $\mathbb{E}(M^-) = -\mathbb{E}(X_1)$.

- Montrer que $\mathbb{E}(e^{itM_n^+}) + \mathbb{E}(e^{-itM_n^-}) - 1 = \mathbb{E}(e^{itM_n}) = \mathbb{E}(e^{itM_{n-1}^+})\mathbb{E}(e^{itX_1})$, pour tous $n \in \mathbb{N}^*$ et $t \in \mathbb{R}$. (Pour la 2ème égalité, on pourra décaler la suite X).
- Déduire le résultat lorsque X_1 est intégrable et $\mathbb{E}(X_1) < 0$. (Vérifier que $M^+ < \infty$ p.s. et que M^- est intégrable).
- Déduire le résultat (par troncation) lorsque $\mathbb{E}(X_1) = -\infty$.

2.2 Temps de retour en 0

Soient $T_0 := 0$, $T_1 = T := \min\{n \in \mathbb{N}^* \mid S_n = 0\}$ le premier temps de retour en 0 (qui vaut ∞ s'il n'y en a pas), puis pour $k \in \mathbb{N}^*$: $T_k := \min\{n > T_{k-1} \mid S_n = 0\}$ le k -ième temps de retour en 0.

Exercice n° 2.2.1 a) Vérifier : $S_{N+k} = S_N + S_k \circ \theta^N$ et $T_k = T_{k-1} + T \circ \theta^{T_{k-1}} = T + T_{k-1} \circ \theta^T$.

- Montrer que $\mathbb{P}(T_k < \infty) = \mathbb{P}(T < \infty)^k$. (Utiliser le lemme 2.)
- Pour tout $v \in \mathbb{Z}^d, v \neq 0$, notons $T_k(v)$ le k -ième temps de passage en v . Montrer que $\mathbb{P}(T_k(v) < \infty) = \mathbb{P}(T_1(v) < \infty) \times \mathbb{P}(T < \infty)^{k-1}$.

Proposition 2 a) Les énoncés suivant sont équivalents :

(i) $\mathbb{P}(T < \infty) = 1$. (ii) $\mathbb{P}(\limsup_n \{S_n = 0\}) = 1$. (iii) $\sum_n \mathbb{P}(S_n = 0) = \infty$.

b) Si $\mathbb{P}(T < \infty) < 1$, alors $\mathbb{P}(\limsup_n \{S_n = v\}) = 0$ pour tout $v \in \mathbb{Z}^d$.

Preuve a) L'exercice 2.2.1.b montre que (i) entraîne $\mathbb{P}(T_n < \infty (\forall n)) = 1$, c'est-à-dire (ii). Le premier lemme de Borel-Cantelli montre que (ii) \Rightarrow (iii). Ensuite nous avons $\sum_n 1_{\{S_n=0\}} = \sum_k 1_{\{T_k < \infty\}}$ presque sûrement, et donc en intégrant : $\sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0) = \sum_{k \in \mathbb{N}} \mathbb{P}(T_k < \infty) = \sum_k \mathbb{P}(T < \infty)^k = \mathbb{P}(T = \infty)^{-1}$ d'après l'exercice 2.2.1.b à nouveau, d'où (iii) \Rightarrow (i).

b) $\mathbb{E} \left[\sum_n 1_{\{S_n=v\}} \right] = \sum_k \mathbb{P}(T_k(v) < \infty) = \mathbb{P}(T_1(v) < \infty) / \mathbb{P}(T = \infty) < \infty$ d'après l'exercice 2.2.1.c, et donc $\mathbb{P}(\limsup_n \{S_n = v\}) = \mathbb{P}[\sum_n 1_{\{S_n=v\}} = \infty] = 0$. \diamond

Proposition 3 Soient $\mathcal{R} := \{v \in \mathbb{Z}^d \mid \limsup_n \{S_n = v\} = \Omega \text{ p.s.}\}$ l'ensemble des points presque sûrement infiniment visités par la marche aléatoire, et $\mathcal{E} := \{v \in \mathbb{Z}^d \mid (\exists n \in \mathbb{N}) \mathbb{P}(S_n = v) > 0\}$ l'ensemble des points pouvant être visités par la marche aléatoire.

Alors soit \mathcal{R} est vide, soit c'est un sous-groupe de \mathbb{Z}^d et il est égal à \mathcal{E} .

Preuve Soient $v \in \mathcal{E}$, et $w \notin v + \mathcal{R}$. Fixons $k \in \mathbb{N}$ tel que $\mathbb{P}(S_k = v) > 0$, et $m \in \mathbb{N}^*$ tel que $\mathbb{P}((\forall n \geq m) S_n \neq w - v) > 0$. Alors

$$\begin{aligned} \mathbb{P}((\forall n \geq m+k) S_n \neq w) &\geq \mathbb{P}((\forall n \geq m+k) S_n - S_k \neq w - v, S_k = v) \\ &= \mathbb{P}((\forall n \geq m+k) S_n - S_k \neq w - v) \mathbb{P}(S_k = v) = \mathbb{P}((\forall n \geq m) S_n \neq w - v) \mathbb{P}(S_k = v) > 0. \end{aligned}$$

Donc nous venons de prouver que si $w - v \notin \mathcal{R}$ et $v \in \mathcal{E}$, alors $w \notin \mathcal{R}$. Autrement dit : $v \in \mathcal{E}$ et $w \in \mathcal{R}$ impliquent $w - v \in \mathcal{R}$.

Remarquons que $\mathcal{R} \subset \mathcal{E}$. Donc prenant v et w dans \mathcal{R} , nous voyons que \mathcal{R} est bien un sous-groupe, s'il est non vide. Enfin, dans ce cas toujours, prenant $w = 0$, nous obtenons $\mathcal{E} \subset \mathcal{R}$. \diamond

Exercice n° 2.2.2 a) Donner un exemple où \mathcal{E} n'est pas un sous-groupe de \mathbb{Z}^d .

b) Montrer que si \mathcal{E} est un sous-groupe de \mathbb{Z}^d , alors c'est le sous-groupe engendré par le support de μ .

Exercice n° 2.2.3 Considérons une probabilité μ sur \mathbb{Z}^d qui ne charge pas 0, un réel $q \in]0, 1[$, et la probabilité μ' sur \mathbb{Z}^d définie par : $\mu'(\{0\}) = 1 - q$ et $\mu'(\{v\}) = q \mu(\{v\})$ pour tout $v \neq 0$. Soient S et S' les deux marches associées.

Soit $\{r_j \mid j \in \mathbb{N}^*\}$ une suite indépendante de S de v.a.i.i.d. géométriques de paramètre q (i.e. $\mathbb{P}(r_j = k) = (1 - q)^{k-1} q$), et soient $R_0 := 0$, puis pour $k \in \mathbb{N}^*$: $R_k := r_1 + \dots + r_k$.

a) Montrer que pour tout $m \in \mathbb{N}^*$, $\mathbb{P}[(\exists k \in \mathbb{N}^*) R_k = m] = q$. (Par récurrence.)

Posons ensuite $S_n'' := S_k$ sur $\{R_k \leq n < R_{k+1}\}$, pour tout $n \in \mathbb{N}$.

b) Montrer que S'' et S' ont la même loi. (Évaluer $\mathbb{P}\left(\bigcap_{\ell=1}^k \bigcap_{m=0}^{j_\ell-1} \{S''_{j_1+\dots+j_{\ell-1}+m} = z_{\ell-1}\}\right)$, pour $j_1, \dots, j_k \in \mathbb{N}^*$ et $0 = z_0 \neq z_1 \neq \dots \neq z_{k-1}$.)

c) Dédurre que $\sum_k \mathbb{P}(S_k = 0) = q \sum_n \mathbb{P}(S'_n = 0)$, et que S est récurrente ssi S' l'est.

d) Comparer $\mathbb{P}(T' < \infty)$ et $\mathbb{P}(T < \infty)$. (T' est le temps de retour en 0 pour S' .)

e) Soient $R' := \min\{n \in \mathbb{N} \mid S'_n \neq 0\}$, $R'_0 := 0$, puis pour $k \in \mathbb{N}^*$: $R'_k := R'_{k-1} + R' \circ \theta^{R'_{k-1}}$. Identifier la loi de la suite R' , puis montrer que la suite $S'_{R'_k}$ a la loi de S . et est indépendante de la suite R' . (Évaluer $\mathbb{P}\left[\bigcap_{\ell=1}^k \{R'_\ell = j_\ell; S'_{R'_\ell} = z_\ell\}\right]$, par récurrence sur k .)

2.3 Récurrence des marches aléatoires simples

Dans le cas de la marche simple sur \mathbb{Z}^d (voir la définition 4), il est clair que $\mathcal{E} = \mathbb{Z}^d$. Les propositions 2 et 3 nous assurent donc que soit la marche simple repasse indéfiniment par chaque point de \mathbb{Z}^d , soit chaque point n'est visité qu'un nombre fini de fois par elle, et donc elle tend vers l'infini, tout ceci presque sûrement. Dans le premier cas, on dit que la marche est récurrente, et dans le second, qu'elle est transitoire (ou transiente).

Nota Bene: Tirages d'urnes, loi multinômiale, formule multinômiale.

Fixons deux entiers $N, d \geq 2$, et une décomposition $N = n_1 + \dots + n_d$ en entiers naturels $n_1, \dots, n_d \in \mathbb{N}$. Considérons une urne remplie de N boules, dont (pour $1 \leq j \leq d$) n_j sont marquées j , dans laquelle on pratique un tirage exhaustif des N boules (une par une et sans remise). On se demande quel est le nombre de tirages possibles. Imaginant un instant les N boules distinguables, un tel tirage est modélisé par une permutation $\sigma \in \mathcal{S}_N$. Mais comme la question ne distingue les boules que par leur numéro (et non les boules portant un même numéro), on cherche en fait le cardinal de l'ensemble $\mathcal{P}_N(n_1, \dots, n_d)$ des partitions

de $\{1, \dots, N\}$ en $\{1, \dots, N\} = \bigsqcup_{j=1}^d A_j$, avec (pour $1 \leq j \leq d$) $\text{Card}(A_j) = n_j$. Notant $\mathcal{S}_N(A_1, \dots, A_d) := \{\sigma \in \mathcal{S}_N \mid \sigma[1, \dots, n_1] = A_1, \dots, \sigma[n_1 + \dots + n_{d-1} + 1, \dots, N] = A_d\}$ l'ensemble des permutations produisant le tirage correspondant à la partition (A_1, \dots, A_d) ,

nous avons d'une part $\mathcal{S}_N = \bigsqcup \left\{ \mathcal{S}_N(A_1, \dots, A_d) \mid (A_1, \dots, A_d) \in \mathcal{P}_N(n_1, \dots, n_d) \right\}$, et d'autre part $\text{Card}[\mathcal{S}_N(A_1, \dots, A_d)] = n_1! \times \dots \times n_d!$. De sorte que le nombre cherché est $\text{Card}[\mathcal{P}_N(n_1, \dots, n_d)] = \frac{N!}{n_1! \times \dots \times n_d!}$. Ceci équivaut à la formule du d -nôme: pour tous

$z_1, \dots, z_d \in \mathbb{C}$, on a $(z_1 + \dots + z_d)^N = \sum_{\substack{n_1, \dots, n_d \in \mathbb{N} \\ n_1 + \dots + n_d = N}} \frac{N!}{n_1! \dots n_d!} z_1^{n_1} \dots z_d^{n_d}$. En effet, il est clair a

priori que $(z_1 + \dots + z_d)^N$ est une telle combinaison linéaire de ces monômes $z_1^{n_1} \dots z_d^{n_d}$. Et ce dernier intervient exactement autant de fois qu'il y a de façons de séparer les N termes $(z_1 + \dots + z_d)$ en: n_1 dans lesquels on choisit z_1 , \dots , n_d dans lesquels on choisit z_d ; donc au total précisément le "coefficient multinômial" $\text{Card}[\mathcal{P}_N(n_1, \dots, n_d)]$. Noter qu'en prenant $z_1 = \dots = z_d = 1$, on a $d^N = \sum_{n_1, \dots, n_d} \frac{N!}{n_1! \dots n_d!}$. De sorte que les $\frac{d^{-N} N!}{n_1! \dots n_d!}$ définissent

sur $\left\{ (n_1, \dots, n_d) \in \mathbb{N}^d \mid n_1 + \dots + n_d = N \right\}$ une loi de probabilité, la "loi multinômiale". Pour $d = 2$, on retrouve bien entendu la formule du binôme et la loi binômiale.

Le célèbre résultat suivant peut se traduire en langage courant par la phrase : un homme ivre finira par retrouver sa demeure, alors qu'un oiseau ivre peut se perdre définitivement.

Théorème 2 (Pólya) *La marche aléatoire simple sur \mathbb{Z}^d est récurrente pour $d = 1$ ou $d = 2$, mais est transitoire pour $d \geq 3$.*

Preuve La marche est récurrente ssi 0 est récurrent, i.e. ssi la marche repasse presque sûrement indéfiniment par 0, et donc, d'après la proposition 2, ssi $\sum_n \mathbb{P}(S_n = 0) = \infty$. Comme $\mathbb{P}(S_{2n+1} = 0) = 0$ pour tout n , nous avons la récurrence ssi $\sum_n \mathbb{P}(S_{2n} = 0) = \infty$.

Pour $d = 1$, nous avons aussitôt $\mathbb{P}(S_{2n} = 0) = 2^{-2n} C_{2n}^n \sim (\pi n)^{-1/2}$, puisqu'il faut autant de pas vers la gauche que vers la droite, et en utilisant la formule de Stirling.

Pour $d = 2$, distinguant selon le nombre m de déplacements de la marche dans le sens horizontal, on retrouve la loi multinômiale : $\mathbb{P}(S_{2n} = 0) = 4^{-2n} \sum_{m=0}^n \frac{(2n)!}{m!m!(n-m)!(n-m)!}$, d'où $\mathbb{P}(S_{2n} = 0) = 2^{-4n} C_{2n}^n \sum_{m=0}^n C_n^m C_n^{n-m} = (2^{-2n} C_{2n}^n)^2 \sim (\pi n)^{-1}$.

Pour $d = 3$, nous avons, de nouveau selon la loi multinômiale :

$$\begin{aligned} \mathbb{P}(S_{2n} = 0) &= 6^{-2n} \sum_{0 \leq l+m \leq n} \frac{(2n)!}{[l!m!(n-l-m)!]^2} = 2^{-2n} C_{2n}^n \sum_{0 \leq l+m \leq n} \left(3^{-n} \frac{n!}{l!m!(n-l-m)!} \right)^2 \\ &\leq 2^{-2n} C_{2n}^n \max_{0 \leq l+m \leq n} \left\{ 3^{-n} \frac{n!}{[l!m!(n-l-m)!]^2} \right\} \sum_{0 \leq l+m \leq n} \left(3^{-n} \frac{n!}{l!m!(n-l-m)!} \right) \\ &= 2^{-2n} C_{2n}^n \max_{n_1+n_2+n_3=n} \left\{ 3^{-n} \frac{n!}{n_1!n_2!n_3!} \right\} = \frac{2^{-2n} C_{2n}^n \times 3^{-n} \times n!}{\min \{ n_1!n_2!n_3! \mid n_1+n_2+n_3=n, |n_i-n_j| \leq 1 \}}, \end{aligned}$$

puisque si par exemple $n_1 - n_2 \geq 2$ alors $n_1!n_2!n_3! > (n_1 - 1)!(n_2 + 1)!n_3!$; cela impose $n_1, n_2, n_3 \geq (n - 2)/3$, d'où par la formule de Stirling la conclusion pour $d = 3$:

$$\mathbb{P}(S_{2n} = 0) = \mathcal{O}\left(\frac{n^{-1/2} \times 3^{-n} \times (n/e)^n \sqrt{n}}{(n_1/e)^{n_1} (n_2/e)^{n_2} (n_3/e)^{n_3} \sqrt{n_1 n_2 n_3}} \right) = \mathcal{O}\left(\frac{3^{-n} \times n^n}{\left(\frac{n-2}{3}\right)^{n+3/2}} \right) = \mathcal{O}(n^{-3/2}).$$

Finalement, si $d > 3$, considérons la projection S'_n de la marche S_n sur les trois premières coordonnées de \mathbb{Z}^d , puis $T_0 := 0$, et pour $n \in \mathbb{N}^*$:

$T_n := \min\{m > T_{n-1} \mid S'_m \neq S'_{T_{n-1}}\}$. Alors T_n est p.s. fini, S'_{T_n} est la marche simple sur \mathbb{Z}^3 , et donc ne passe plus par 0 après un certain temps aléatoire presque sûrement fini N . Comme S'_n est constante entre deux instants T_{n-1} et T_n , nous avons a fortiori $S'_n \neq 0$ (et donc $S_n \neq 0$) pour $n > T_N$, d'où le caractère transitoire annoncé. \diamond

Exercice n° 2.3.1 Marche aléatoire simple sur un arbre homogène (connexe).

Soit \mathcal{A} un arbre : c'est un graphe (non vide) non orienté sans boucle, de sorte que 2 points (ou sommets) quelconques de l'arbre sont reliés par un unique chemin injectif ; le nombre d'arêtes que comporte ce chemin définit la distance entre ces deux points ; elle fait de

l'arbre un espace métrique. Supposons \mathcal{A} homogène : chaque point admet exactement d voisins (i.e. points situés à distance 1). À quoi \mathcal{A} est-il isométrique lorsque $d = 1$ ou 2 ? Fixons $d \geq 3$, et un point $O \in \mathcal{A}$. La marche simple S sur \mathcal{A} part de $S_0 := O$, et pour tout $n \in \mathbb{N}^*$, la loi de S_n sachant S_{n-1} est uniforme sur l'ensemble des d voisins de S_{n-1} . Notons δ_n la distance entre S_0 et S_n .

a) Montrer qu'il existe une suite b de v.a.i.i.d., de loi à déterminer, telle que pour tout $n \in \mathbb{N}$: $\delta_{n+1} - \delta_n = b_n 1_{\{\delta_n > 0\}} + 1_{\{\delta_n = 0\}}$. (Définir b_n sur $\{\delta_n = 0\}$ par b'_n indépendante ayant la bonne loi, puis vérifier que b convient en calculant $\mathbb{P}(b_n = 1 | b'_{n-1}, \dots, b'_1, \mathcal{F}_n)$.)

b) Montrer que la marche S est transitoire, et qu'elle admet une vitesse de fuite, limite presque sûre de δ_n/n , à déterminer. (Commencer par minorer δ_n .)

2.4 Récurrence des autres marches aléatoires

En général, la marche aléatoire est dite récurrente lorsque \mathcal{R} est non vide, et transitoire (ou transiente, d'après l'anglais) sinon.

Les propositions 2 et 3 entraînent qu'il y a récurrence ssi $\sum_n \mathbb{P}(S_n = 0) = \infty$.

Nous supposons que la loi μ du déplacement de la marche admet un premier moment.

Fixons pour norme sur \mathbb{Z}^d le maximum des valeurs absolues des coordonnées. Nous aurons besoin du lemme suivant.

Lemme 3 *Pour tout $M \in \mathbb{N}$, nous avons :*

$$\sum_{n \in \mathbb{N}} \mathbb{P}(\|S_n\| \leq M) \leq (2M + 1)^d \sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0).$$

Preuve Posons $E := \{-M, \dots, M\}^d$, et pour tout $v \in E$, $T_v := \min\{n \in \mathbb{N} | S_n = v\}$. Nous avons :

$$\begin{aligned} \sum_{n \in \mathbb{N}} \mathbb{P}(\|S_n\| \leq M) &= \sum_{n \in \mathbb{N}} \sum_{v \in E} \mathbb{P}(S_n = v) = \sum_{v \in E} \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} \mathbb{P}(S_n = v, T_v = m) \\ &= \sum_{v \in E} \sum_{m \in \mathbb{N}} \sum_{n \geq m} \mathbb{P}(S_n - S_m = 0, T_v = m) = \sum_{v \in E} \sum_{m \in \mathbb{N}} \sum_{n \geq m} \mathbb{P}(S_n - S_m = 0) \mathbb{P}(T_v = m) \\ &= \sum_{v \in E} \sum_{l \in \mathbb{N}} \mathbb{P}(S_l = 0) \sum_{m \in \mathbb{N}} \mathbb{P}(T_v = m) \leq \text{Card}(E) \sum_{l \in \mathbb{N}} \mathbb{P}(S_l = 0). \quad \diamond \end{aligned}$$

Voici un résultat de Chung et Fuchs pour le cas de la dimension 1.

Théorème 3 *Une marche aléatoire (à pas intégrable) sur \mathbb{Z} est récurrente si et seulement si elle est centrée.*

Preuve La loi des grands nombres nous dit que S_n/n converge vers $\mathbb{E}(X_1)$. Donc si $\mathbb{E}(X_1) \neq 0$, $S_n \rightarrow \pm\infty$, et donc la marche est transitoire. Supposons ensuite que $\mathbb{E}(X_1) = 0$, et donc que $S_n = o(n)$. Le lemme 3 entraîne :

$$\sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0) \geq (2M + 1)^{-1} \sum_{n \in \mathbb{N}} \mathbb{P}(\|S_n\| \leq M) \geq (2M + 1)^{-1} \sum_{n=0}^{AM} \mathbb{P}(\|S_n\| \leq n/A),$$

pour $M \in \mathbb{N}$ et $A \in \mathbb{N}^*$. Puisque $\lim_n \mathbb{P}(\|S_n\| \leq n/A) = 1$, le critère de Césaro nous donne, en faisant tendre M vers l'infini : $\sum_n \mathbb{P}(S_n = 0) \geq A/2$, et donc $\sum_n \mathbb{P}(S_n = 0) = \infty$, puisque A est arbitraire. \diamond

Voici le résultat analogue pour le cas de la dimension 2.

Théorème 4 *Si μ admet un second moment, la marche aléatoire sur \mathbb{Z}^2 est récurrente si et seulement si elle est centrée.*

Preuve Comme pour $d = 1$, on voit que $\|S_n\| \rightarrow \infty$ si la marche n'est pas centrée. Sinon, le théorème limite central assure que S_n/\sqrt{n} converge en loi vers une gaussienne centrée, de matrice de covariance $K = K(X_1)$.

Si $K = 0$, alors $\mathbb{E}(\|X_1\|^2) = 0$, et la marche reste presque sûrement constamment en 0.

Si K est de rang 1, elle admet deux vecteurs propres orthogonaux u, v tels que $\mathbb{E}[(u \cdot X_1)^2] = 0$ et $\mathbb{E}[(v \cdot X_1)^2] = 1$. Alors la marche reste presque sûrement sur la droite dirigée par v , et nous sommes ramenés au cas $d = 1$ du théorème 3.

Supposons donc que K est inversible. La loi gaussienne limite admet alors une densité f continue > 0 sur \mathbb{R}^2 . Par ailleurs le lemme 3 entraîne :

$$\sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0) \geq (2M + 1)^{-2} \sum_{n \in \mathbb{N}} \mathbb{P}(\|S_n\| \leq M) = \frac{M^2}{(2M + 1)^2} \int_0^\infty \mathbb{P}(\|S_{[sM^2]}\| \leq M) ds,$$

d'où par le théorème limite central et le lemme de Fatou :

$$\sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0) \geq \frac{1}{4} \int_A^\infty \int_{[-s^{-1/2}, s^{-1/2}]^2} f(x) dx ds \geq \frac{1}{4} \int_A^\infty \frac{2f(0)}{s} ds = \infty,$$

où A est tel que $f \geq f(0)/2$ sur $[-A^{-1/2}, A^{-1/2}]^2$. \diamond

2.5 Retours en 0 et loi Arcsinus pour la marche simple sur \mathbb{Z}

Concentrons-nous ici sur la marche simple sur \mathbb{Z} . Nous allons nous intéresser au temps T de son premier retour en 0, puis au temps qu'elle passe dans \mathbb{N} .

Exercice n° 2.5.1 Soit $n \in \mathbb{N}$.

a) Montrer que si $d = 1$ alors pour tout $v \in \mathbb{Z}$: $\mathbb{P}(S_n = v) = 1_{2\mathbb{N}}(n - |v|) \times 2^{-n} C_n^{\frac{n+v}{2}}$.

b) Montrer que dans le cas général, pour $v = (v_1, \dots, v_d) \in \mathbb{Z}^d$:

$$\mathbb{P}(S_n = v) = (2d)^{-n} \times \sum_{n_1 + \dots + n_d = n} \frac{n!}{n_1! \dots n_d!} \prod_{j=1}^d 1_{2\mathbb{N}}(n_j - |v_j|) \times C_{n_j}^{\frac{n_j + v_j}{2}}.$$

Le résultat suivant est connu sous le nom de principe de réflexion de Désiré André.

Lemme 4 Pour tous $n, m, a, b \in \mathbb{N}$, nous avons :

$$\mathbb{P}(S_n = a; S_{n+m} = b; (\exists k \in [n, n+m]) S_k = 0) = \mathbb{P}(S_n = -a; S_{n+m} = b).$$

Preuve Notons $T := \min\{k \geq n \mid S_k = 0\}$. Considérons l'ensemble E des trajectoires de la marche (restreinte à $[n, n+m]$) dont le graphe relie le point (n, a) au point $(n+m, b)$ et rencontre l'axe horizontal, puis l'ensemble E' des trajectoires de la marche (restreinte à $[n, n+m]$) dont le graphe relie le point $(n, -a)$ au point $(n+m, b)$. Considérons aussi l'application qui va de E dans E' , associant à ω la trajectoire ω' déduite de ω en remplaçant la partie du graphe de ω située avant $T(\omega)$ par son symétrique (relativement à l'axe horizontal). C'est une bijection de E sur E' , et donc $\text{Card}(E) = \text{Card}(E')$. Le résultat en découle aussitôt. \diamond

Corollaire 2 Un scrutin donne a voix à un candidat A et b voix à un candidat B , avec $a > b$. La probabilité qu'au cours du dépouillement le score du candidat A dépasse toujours strictement le score du candidat B est $\frac{a-b}{a+b}$.

Preuve Notons S_n le score de A - le score de B après le n -ième bulletin. C'est une marche aléatoire simple sur \mathbb{Z} , et nous cherchons la probabilité qu'elle reste > 0 sur $[1, a+b]$, sachant que $S_{a+b} = a-b$. Utilisant le lemme 4 et l'exercice 2.5.1.a, nous avons :

$$\begin{aligned} \mathbb{P}((\forall k \in [1, a+b]) S_k > 0 \mid S_{a+b} = a-b) &= \frac{\mathbb{P}(S_1 = 1; (\forall k \in [1, a+b]) S_k \neq 0; S_{a+b} = a-b)}{\mathbb{P}(S_{a+b} = a-b)} \\ &= \frac{\mathbb{P}(S_1 = 1; S_{a+b} = a-b) - \mathbb{P}(S_1 = -1; S_{a+b} = a-b)}{\mathbb{P}(S_{a+b} = a-b)} \\ &= \frac{\mathbb{P}(S_1 = 1)\mathbb{P}(S_{a+b-1} = a-b-1) - \mathbb{P}(S_1 = -1)\mathbb{P}(S_{a+b-1} = a-b+1)}{\mathbb{P}(S_{a+b} = a-b)} \\ &= \frac{2^{-a-b}(C_{a+b-1}^{a-1} - C_{a+b-1}^a)}{2^{-a-b}C_{a+b}^a} = \frac{a-b}{a+b}. \quad \diamond \end{aligned}$$

Corollaire 3 $\mathbb{P}(T > 2n) = \mathbb{P}(S_1 \neq 0, \dots, S_{2n} \neq 0) = \mathbb{P}(S_{2n} = 0) \sim (\pi n)^{-1/2}$.

Preuve Utilisant la preuve du corollaire 2 ci-dessus, nous avons :

$$\begin{aligned} \mathbb{P}(S_1 \neq 0, \dots, S_{2n} \neq 0) &= 2\mathbb{P}(S_1 > 0, \dots, S_{2n} > 0) = 2 \sum_{m \in \mathbb{N}^*} \mathbb{P}(S_1 > 0, \dots, S_{2n-1} > 0, S_{2n} = 2m) \\ &= \sum_{m \in \mathbb{N}^*} \left[\mathbb{P}(S_{2n-1} = 2m - 1) - \mathbb{P}(S_{2n-1} = 2m + 1) \right] = \mathbb{P}(S_{2n-1} = 1) \\ &= 2^{1-2n} C_{2n-1}^m = 2^{-2n} C_{2n}^m = \mathbb{P}(S_{2n} = 0). \quad \diamond \end{aligned}$$

Nous déduisons de là le comportement asymptotique de T_n , qui est un exemple de convergence en loi différente du théorème central limite.

Proposition 4 La suite T_n/n^2 converge en loi, vers une v.a. σ telle que

$$\mathbb{E}(e^{-s\sigma}) = e^{-\sqrt{2s}} \quad \text{pour tout } s \geq 0.$$

Preuve Nous allons montrer que les transformées de Laplace de ces lois convergent simplement sur \mathbb{R}_+ , vers une fonction continue en 0. Un théorème de Paul Lévy assure que cela entraîne bien la convergence en loi.

Le lemme 2 et l'exercice 2.2.1 assurent que les v.a. $T_n - T_{n-1}$ forment une suite i.i.d. ; nous avons donc, pour $s > 0$ fixé : $\mathbb{E}(e^{-sT_n/n^2}) = \left(\mathbb{E}(e^{-sT/n^2}) \right)^n$. Or

$$\begin{aligned} \mathbb{E}(e^{-sT/n^2}) &= \sum_{k \in \mathbb{N}^*} e^{-2sk/n^2} \mathbb{P}(T = 2k) = \sum_{k \in \mathbb{N}^*} e^{-2sk/n^2} \left(\mathbb{P}(T > 2k - 2) - \mathbb{P}(T > 2k) \right) \\ &= e^{-2s/n^2} + \sum_{k \in \mathbb{N}^*} \left(e^{-2s(k+1)/n^2} - e^{-2sk/n^2} \right) \mathbb{P}(T > 2k) = e^{-2s/n^2} - \frac{2s}{n^2} (1 + \varepsilon_n) \sum_{k \in \mathbb{N}^*} \frac{1 + \varepsilon'_k}{\sqrt{\pi k}} e^{-2sk/n^2}, \end{aligned}$$

où $\lim_n \varepsilon_n = \lim_k \varepsilon'_k = 0$. Ensuite, comparant la série ci-dessus avec une intégrale, nous avons :

$$\sum_{k \in \mathbb{N}^*} \frac{(1 + \varepsilon'_k)}{\sqrt{\pi k}} e^{-2sk/n^2} \sim n / \sqrt{2s},$$

puisque pour $a \downarrow 0$:

$$\sum_{k \in \mathbb{N}^*} \frac{(1 + \varepsilon'_k)}{\sqrt{k}} e^{-ak} = \int_1^\infty \frac{(1 + \varepsilon'_t)}{\sqrt{t}} e^{-at} dt + \mathcal{O}(1) = \frac{1}{\sqrt{a}} \int_a^\infty \frac{(1 + \varepsilon'_{t/a})}{\sqrt{t}} e^{-t} dt + \mathcal{O}(1) \sim \sqrt{\pi/a}.$$

Nous avons finalement :

$$\mathbb{E}(e^{-sT_n/n^2}) = \left(e^{-2s/n^2} - \frac{\sqrt{2s}}{n} (1 + \varepsilon_n) \right)^n = \left(1 - \frac{\sqrt{2s}}{n} (1 + \varepsilon_n) \right)^n \longrightarrow e^{-\sqrt{2s}}. \quad \diamond$$

Exercice n° 2.5.2 Montrer que T et σ ne sont pas intégrables, mais que σ est p.s. finie > 0 et que $\mathbb{E}(\sigma^{-r}) < \infty$ pour tout $r \geq 0$. (Utiliser $\Gamma(r)$ et un changement de variable).

Proposition 5 Soit $L_{2n} := \max\{m \leq 2n \mid S_m = 0\}$, pour tout $n \in \mathbb{N}$. Alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(L_{2n} \leq 2na) = \frac{2}{\pi} \operatorname{Arcsin}(\sqrt{a}), \quad \text{pour tout } a \in [0, 1].$$

Preuve Nous déduisons d'abord du corollaire 3 que pour tout $k \leq n$:

$$\mathbb{P}(L_{2n} = 2k) = \mathbb{P}(S_{2k} = 0, S_{2k+1} \neq 0, \dots, S_{2n} \neq 0) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0).$$

Ensuite, pour $0 < \varepsilon < a < 1$ fixés, nous avons :

$$\begin{aligned} \mathbb{P}(2n\varepsilon < L_{2n} \leq 2na) &= \sum_{k=[n\varepsilon]+1}^{[na]} \mathbb{P}(L_{2n} = 2k) = \int_{[n\varepsilon]+1}^{[na]+1} \mathbb{P}(L_{2n} = 2[s]) ds \\ &= \int_{\frac{[n\varepsilon]+1}{n}}^{\frac{[na]+1}{n}} n \mathbb{P}(L_{2n} = 2[ns]) ds = \int_{\frac{[n\varepsilon]+1}{n}}^{\frac{[na]+1}{n}} \sqrt{n} 2^{-2[ns]} C_{2[ns]}^{[ns]} \times \sqrt{n} 2^{2[ns]-2n} C_{2n-2[ns]}^{n-[ns]} ds \\ &\rightarrow \int_{\varepsilon}^a (\pi s)^{-1/2} (\pi(1-s))^{-1/2} ds = \frac{2}{\pi} \int_{\sqrt{\varepsilon}}^{\sqrt{a}} \frac{ds}{\sqrt{1-s^2}} = \frac{2}{\pi} [\operatorname{Arcsin}(\sqrt{a}) - \operatorname{Arcsin}(\sqrt{\varepsilon})], \end{aligned}$$

car l'équivalent employé dans l'intégrale est uniforme sur $[\varepsilon, (a+1)/2]$. On passe à $\varepsilon = 0$ en utilisant ci-dessus que $\sqrt{ns} 2^{-2[ns]} C_{2[ns]}^{[ns]}$ est borné sur $\{n \in \mathbb{N}, s \in \mathbb{R}_+\}$, d'où

$$\mathbb{P}\left[\frac{L_{2n}}{2n} \leq a\right] = \mathcal{O}\left(\int_0^a \frac{ds}{\sqrt{\pi(1-s)s}}\right) = \mathcal{O}(\sqrt{a}), \quad \text{qu'on applique à } \varepsilon \text{ au lieu de } a. \quad \diamond$$

Remarque 1 Si deux joueurs jouent à pile ou face chaque jour de l'année civile, il y a environ une chance sur deux qu'un des deux joueurs soit en tête du jeu du 1-7 au 31-12 !

Proposition 6 Soit $\Lambda_{2n} := \operatorname{Card}\{m < 2n \mid S_m \vee S_{m+1} > 0\}$, pour tout $n \in \mathbb{N}$. Alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(\Lambda_{2n} \leq 2na) = \frac{2}{\pi} \operatorname{Arcsin}(\sqrt{a}), \quad \text{pour tout } a \in [0, 1].$$

Preuve Il suffit d'utiliser la proposition précédente, en prouvant que Λ_{2n} et L_{2n} ont la même loi, i.e. $\mathbb{P}(\Lambda_{2n} = 2k) = \mathbb{P}(L_{2n} = 2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0)$, pour $n \in \mathbb{N}$ et $0 \leq k \leq n$. Faisons ceci par récurrence sur n . C'est clair pour $0 \leq k \leq n \leq 1$. Supposons l'égalité vraie pour les (k, n') tels que $0 \leq k \leq n' < n$.

Pour $k = n$, nous déduisons du corollaire 3 :

$$\begin{aligned} \mathbb{P}(\Lambda_{2n} = 2n) &= \mathbb{P}(S_1 \geq 0, \dots, S_{2n} \geq 0) = \mathbb{P}(S_1 \geq 0, \dots, S_{2n-1} \geq 0) \\ &= \mathbb{P}(S_2 - S_1 \geq 0, \dots, S_{2n} - S_1 \geq 0) = 2 \mathbb{P}(S_1 = 1, S_2 - S_1 \geq 0, \dots, S_{2n} - S_1 \geq 0) \\ &= 2 \mathbb{P}(S_1 > 0, \dots, S_{2n} > 0) = \mathbb{P}(S_{2n} = 0) = \mathbb{P}(L_{2n} = 2n). \end{aligned}$$

Changeant S en $-S$, i.e. X en $-X$, nous voyons que le résultat pour $k = n$ est équivalent au résultat pour $k = 0$. Reste à considérer le cas $0 < k < n$. Alors $2 \leq T \leq 2n - 2$. Distinguant suivant la valeur de T et selon que $S_1 = 1$ ou $S_1 = -1$, puis appliquant l'hypothèse de récurrence, nous obtenons :

$$\begin{aligned} \mathbb{P}(\Lambda_{2n} = 2k) &= \frac{1}{2} \sum_{m=1}^k \mathbb{P}(T=2m) \mathbb{P}(\Lambda_{2n-2m} = 2k - 2m) + \frac{1}{2} \sum_{m=1}^{n-k} \mathbb{P}(T=2m) \mathbb{P}(\Lambda_{2n-2m} = 2k) = \\ &= \frac{1}{2} \mathbb{P}(S_{2n-2k} = 0) \sum_{m=1}^k \mathbb{P}(T=2m) \mathbb{P}(S_{2k-2m} = 0) + \frac{1}{2} \mathbb{P}(S_{2k} = 0) \sum_{m=1}^{n-k} \mathbb{P}(T=2m) \mathbb{P}(S_{2n-2k-2m} = 0) \\ &= \frac{1}{2} \mathbb{P}(S_{2n-2k} = 0) \mathbb{P}(S_{2k} = 0) + \frac{1}{2} \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0) = \mathbb{P}(L_{2n} = 2k) . \diamond \end{aligned}$$

Exercice n° 2.5.3 Soient $N_n := \sum_{k=1}^n 1_{\{S_k=0\}}$ le nombre de passages en 0 au cours des n premiers pas, $\Lambda'_{2n} := \text{Card} \{m \leq 2n \mid S_m > 0\}$ et $\Lambda''_{2n} := \text{Card} \{m \leq 2n \mid S_m \geq 0\}$, pour tout $n \in \mathbb{N}$.

- a) Dédurre de la proposition 4 que n/T_n et N_n/n convergent en probabilité vers 0.
b) Dédurre que Λ'_{2n} et Λ''_{2n} obéissent à la même loi limite de l'Arcsinus que Λ_{2n} .

3 Chaînes de Markov

3.1 La propriété de Markov

Définition 6 Soit E un ensemble dénombrable (qui peut être fini), dit espace des états. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une suite $\{X_n \mid n \in \mathbb{N}\}$ de v.a. de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{P}(E))$ est une chaîne de Markov sur E lorsque pour tout $n \in \mathbb{N}$

la loi de X_{n+1} sachant X_0, \dots, X_n est la même que la loi de X_{n+1} sachant X_n , id est :

$$\mathbb{P}(X_{n+1} = e_{n+1} \mid X_0 = e_0, \dots, X_n = e_n) = \mathbb{P}(X_{n+1} = e_{n+1} \mid X_n = e_n), \forall e_0, \dots, e_{n+1} \in E.$$

La propriété de Markov, écrite précisément ci-dessus, signifie que la loi du futur ne dépend que du présent, et non du passé : il n'y a pas d'inertie.

Exemple Toute marche aléatoire (sur \mathbb{Z}^d en particulier) est une chaîne de Markov :

$$\begin{aligned} \mathbb{P}(S_{n+1} = v \mid S_1 = v_1, \dots, S_n = v_n) &= \mathbb{P}(S_{n+1} - S_n = v - v_n \mid S_1 = v_1, \dots, S_n = v_n) \\ &= \mathbb{P}(S_{n+1} - S_n = v - v_n) = \mu(\{v - v_n\}) = \mathbb{P}(S_{n+1} - S_n = v - v_n \mid S_n = v_n) \\ &= \mathbb{P}(S_{n+1} = v \mid S_n = v_n). \end{aligned}$$

Exercice n° 3.1.0 Vérifier que le processus X est une chaîne de Markov si et seulement si pour tout $n \in \mathbb{N}$ il existe une fonction φ_n de E^2 dans $[0, 1]$ telle que pour tous $e_0, \dots, e_{n+1} \in E$,

$$\mathbb{P}(X_{n+1} = e_{n+1} \mid X_0 = e_0, \dots, X_n = e_n) = \varphi_n(e_n, e_{n+1}).$$

Rappel Pour tout $n \in \mathbb{N}$, notons $\mathcal{F}_n := \sigma(\{X_0, \dots, X_n\})$ la sous-tribu de \mathcal{F} qui représente l'information connue au temps n .

La notion de temps d'arrêt est la même pour les chaînes de Markov que pour les marches aléatoires : revoir la définition 5. De même pour la tribu \mathcal{F}_T , et pour les lois \mathbb{P}_x et \mathbb{P}_ν . Pour que X_T ait un sens même si T peut être infini, on pose $X_\infty = \delta$, où $\delta \notin E$ est un point isolé ajouté à E , appelé cimetière.

Définition 7 On appelle probabilité de transition ou noyau de transition sur l'espace dénombrable E toute application P de $E \times E$ dans $[0, 1]$ telle que pour tout $e \in E$, $P(e, \cdot)$ définit une probabilité sur $\mathcal{P}(E)$: $P(e, e') \geq 0$ ($\forall e' \in E$), et $\sum_{e' \in E} P(e, e') = 1$.

Si $\{X_n \mid n \in \mathbb{N}\}$ est une chaîne de Markov de $(\Omega, \mathcal{F}, \mathbb{P})$ dans E , ses probabilités de transition sont par définition les noyaux P_n tels que $P_n(X_n, A) = \mathbb{P}(X_{n+1} \in A \mid X_n)$, id est tels que $P_n(e, e') = \mathbb{P}(X_{n+1} = e' \mid X_n = e)$ pour tous $e, e' \in E$ et $n \in \mathbb{N}$.

A toute chaîne de Markov correspond bien une suite de probabilités de transition, et réciproquement, à toute suite de probabilités de transition P_n associée à une probabilité μ sur E , correspond par le théorème de Kolmogorov une unique loi de chaîne de Markov X définie sur $\Omega = E^{\mathbb{N}}$ et à valeurs dans E , admettant la suite P_n comme suite de probabilités de transition et μ comme loi initiale. Cette loi \mathbb{P} est déterminée par : pour tout $n \in \mathbb{N}$

$$\mathbb{P}(X_0 = e_0, \dots, X_n = e_n) = \mu(e_0)P_0(e_0, e_1) \dots P_{n-1}(e_{n-1}, e_n), \quad \forall n \in \mathbb{N}, \forall e_0, \dots, e_n \in E.$$

L'ensemble $E^{\mathbb{N}}$ est l'espace des trajectoires possibles pour la chaîne sur E , appelé aussi l'espace canonique, qu'on peut donc toujours prendre comme espace Ω sur lequel la chaîne et sa loi sont définies. Vue comme variable aléatoire sur cet espace canonique, la variable X_n est simplement la n -ième projection canonique, id est la n -ième coordonnée.

La formule ci-dessus définit bien (par projection sur chaque tribu \mathcal{F}_n) une unique probabilité sur l'espace canonique (c'est dû au théorème de Kolmogorov), que l'on note \mathbb{P}_μ . Lorsque μ est la masse de Dirac en un point $e \in E$, on note simplement \mathbb{P}_e : c'est la loi de la chaîne issue de e . Notons qu'il suffit de connaître la famille $\{\mathbb{P}_e \mid e \in E\}$, puisqu'on a simplement $\mathbb{P}_\mu = \sum_{e \in E} \mu(e)\mathbb{P}_e$.

Exercice n° 3.1.1 Vérifier que le processus X défini ci-dessus par ses marginales finidimensionnelles constitue effectivement une chaîne de Markov, et que, réciproquement, si X est une chaîne de Markov de loi initiale μ et de probabilités de transition P_n , alors

les lois marginales fini-dimensionnelles de X sont nécessairement données par la formule ci-dessus. Vérifier qu'on a plus généralement, pour $n \in \mathbb{N}$ et f_0, \dots, f_n fonctions ≥ 0 sur E :

$$\mathbb{E}\left(\prod_{j=0}^n f_j \circ X_j\right) = \sum_{e_0} \mu(e_0) f_0(e_0) \sum_{e_1} P_0(e_0, e_1) f_1(e_1) \sum_{e_2} \dots \sum_{e_n} P_{n-1}(e_{n-1}, e_n) f_n(e_n).$$

Définition 8 Une chaîne de Markov est dite homogène lorsque la suite de ses probabilités de transition est constante.

Si (X_n) est une chaîne de Markov sur E , alors (n, X_n) est une chaîne de Markov homogène sur $\mathbb{N} \times E$. Nous nous limiterons dorénavant aux chaînes de Markov homogènes.

La probabilité de transition sera notée P . Pour une marche, $P(x, y) = \mu(\{y - x\})$. Dans le cas où E est fini de cardinal n , P est ainsi simplement une matrice carrée de format (n, n) "stochastique", i.e. à coefficients dans $[0, 1]$ et telle que la somme de chacune de ses lignes vaut 1. Le cas de E dénombrable est analogue, sauf que la "matrice" de transition P a cette fois une infinité de lignes et de colonnes.

Il est de même très commode d'identifier toute probabilité (et plus généralement toute mesure positive) sur E avec le vecteur-ligne donnant sa loi, de sorte que la coordonnée d'indice $e \in E$ de ce vecteur est la probabilité (ou la mesure) de $\{e\}$.

Sur $\Omega = E^{\mathbb{N}}$ l'opérateur de décalage θ est défini par $X_j \circ \theta^n = X_{j+n}$ ($\forall j, n \in \mathbb{N}$). (Noter la différence avec la section 2.1, si S est la chaîne qu'on considère, au lieu de X).

La propriété de Markov s'enrichit comme suit :

Exercice n° 3.1.2

a) *Propriété de Markov, 2ème état* : Vérifier que la propriété de Markov équivaut à :

Pour $n \in \mathbb{N}, k \in \mathbb{N}^*, x_0, \dots, x_n, y_1, \dots, y_k \in E$, et toute loi ν sur E , nous avons

$$\begin{aligned} \mathbb{P}_\nu(X_{n+k} = y_k, \dots, X_{n+1} = y_1, X_n = x_n, \dots, X_0 = x_0) \\ = \mathbb{P}_{x_n}(X_k = y_k, \dots, X_1 = y_1) \times \mathbb{P}_\nu(X_n = x_n, \dots, X_0 = x_0). \end{aligned}$$

b) *Propriété de Markov, 3ème état* : Dédire que la propriété de Markov équivaut à :

Pour tous $n \in \mathbb{N}, x \in E, B_n \in \mathcal{F}_n$ tel que $B_n \subset \{X_n = x\}$, $A \in \mathcal{F}_\infty$, et toute loi ν sur E , nous avons

$$\mathbb{P}_\nu([\theta^n]^{-1}(A) \cap B_n) = \mathbb{P}_x(A) \times \mathbb{P}_\nu(B_n).$$

c) *Propriété de Markov, 4ème état* : Dédire que la propriété de Markov équivaut à :

Pour tous temps d'arrêt $T, x \in E, B \in \mathcal{F}_T$ tel que $B \subset \{X_T = x\}$, et $A \in \mathcal{F}_\infty$, nous avons

$$\mathbb{P}_\nu([\theta^T]^{-1}(A) \cap B) = \mathbb{P}_x(A) \times \mathbb{P}_\nu(B).$$

d) *Propriété de Markov, 5ème état* : Dédire que la propriété de Markov équivaut à :

Pour toute $Y \geq 0$ \mathcal{F}_∞ -mesurable, tout temps d'arrêt T , et toute loi ν sur E , nous avons

$$\mathbb{E}_\nu(Y \circ \theta^T | \mathcal{F}_T) = \mathbb{E}_{X_T}(Y) \quad \mathbb{P}_\nu - p.s. \text{ sur } \{T < \infty\}.$$

e) *Propriété de Markov, 6ème état* : Dédurre que la propriété de Markov équivaut à :

Pour toute $Y \geq 0$ \mathcal{F}_∞ -mesurable, tout temps d'arrêt T , et toute loi ν sur E , nous avons

$$\mathbb{E}_\nu(Y \circ \theta^T | \mathcal{F}_T) = \mathbb{E}_\nu(Y \circ \theta^T | X_T) \quad \mathbb{P}_\nu - p.s. \text{ sur } \{T < \infty\}.$$

L'exercice 3.1.2 prouve en particulier la très importante propriété de Markov forte :

Proposition 7 (*Propriété de Markov forte*). Soient X une chaîne de Markov, \mathbb{P}_μ sa loi, $x \in E$, et T un temps d'arrêt.

Alors, conditionnellement à l'événement $\{X_T = x\}$, la chaîne $\{X_{T+n} | n \in \mathbb{N}\}$ est markovienne de loi \mathbb{P}_x , et indépendante de \mathcal{F}_T .

Autrement dit, pour toute v.a. $Z \geq 0$ sur Ω : $\mathbb{E}(Z \circ \theta^T | \mathcal{F}_T) = \mathbb{E}_{X_T}(Z)$ sur $\{T < \infty\}$.

Exercice n° 3.1.3 a) Vérifier que si P, Q sont 2 matrices stochastiques, et si $n \in \mathbb{N}$, alors $P \times Q$ et P^n sont aussi des matrices stochastiques.

Vérifier que si la loi initiale de X est μ , alors pour tout n la loi de X_n est μP^n , et, pour tous $m \in \mathbb{N}, x, y \in E$, $\mathbb{P}(X_{n+m} = y | X_m = x) = \mathbb{P}_x(X_n = y) = P^n(x, y)$.

Exemples 1) $E = \mathbb{N}$ et $P(k, \cdot)$ est la loi de S_k , S étant la marche aléatoire associée à une loi ν donnée sur \mathbb{N} . Ici X_n modélise une population à la n -ième génération, chaque individu se reproduisant avant de mourir avec la loi fixe ν , indépendamment des autres. Voir l'exercice 3.1.8.

2) $E = \{0, 1, \dots, n\}$ et $P(k, k+1) = (n-k)/n$, $P(k, k-1) = k/n$. Cette chaîne modélise le nombre de molécules de gaz situées dans une urne (d'Ehrenfest) U à chaque instant, sur un nombre total n de molécules se trouvant soit dans U soit dans une autre urne reliée à U par un mince passage ; la transition d'un instant au suivant s'effectue en tirant une molécule au hasard parmi les n et en la changeant d'urne.

3) $E = \mathbb{N}$ et $P(k, k-1) = 1$ si $k \in \mathbb{N}^*$, $P(0, j) = q_j$ pour $j \in \mathbb{N}$, une probabilité q étant donnée sur \mathbb{N} . La chaîne décroît régulièrement jusqu'à retomber en 0, d'où elle rebondit à chaque fois à une hauteur aléatoire (gouvernée par q), indépendamment d'une fois à l'autre.

On peut visualiser une chaîne de Markov au moyen de son graphe de transition : c'est le graphe orienté dont les sommets (ou nœuds) sont les états (points de E), et les arêtes orientées les couples $(x, y) \in E^2$ tels que $P(x, y) > 0$, sur lesquelles on porte la valeur $P(x, y)$.

Exemples Représenter le graphe de transition des chaînes des exemples 2) et 3) ci-dessus, ainsi que de la marche simple sur \mathbb{Z} ou \mathbb{Z}^2 .

Exercice n° 3.1.4 Fixons une v.a. X_0 à valeurs dans E , une suite $\{Y_n | n \in \mathbb{N}\}$ de v.a.i.i.d. uniformes sur $[0, 1]$, indépendantes de X_0 , une fonction mesurable f de $E \times [0, 1]$ dans E , et posons pour tout $n \in \mathbb{N}$: $X_{n+1} := f(X_n, Y_{n+1})$.

- Montrer que X est une chaîne de Markov homogène, et calculer sa matrice de transition.
- Montrer que toute chaîne de Markov homogène est de ce type. (Comment simuler une loi discrète au moyen d'une loi uniforme ?). Comment simuler par ordinateur une chaîne de Markov homogène ?

Exercice n° 3.1.5 Soient $\{B_n | n \in \mathbb{N}^*\}$ une suite i.i.d. de v.a. de Bernoulli $\mathcal{B}(0, 1; p)$, puis S la marche associée. Déterminer dans chacun des cas suivant si la suite X est markovienne, précisant son espace d'états et sa matrice de transition si oui, et un contre-exemple sinon :

- $X_n = B_n$;
 - $X_n = S_n$;
 - $X_n = S_0 + S_1 + \dots + S_n$;
 - $X_n = (S_n, S_0 + S_1 + \dots + S_n)$.
- e) Montrer qu'en général l'image d'une chaîne de Markov par une application f n'est pas markovienne. Donner une condition suffisante sur f pour que cela soit cependant le cas.

Exercice n° 3.1.6 Une puce saute chaque seconde d'un sommet d'un triangle à un autre, indépendamment et au hasard. Tandis qu'une tique choisit 2 fois plus souvent de sauter dans le sens direct que dans le sens indirect. Dans les deux cas, calculer la probabilité que l'animal se retrouve à son sommet de départ au bout de n secondes.

Exercice n° 3.1.7 Soient T_1, \dots, T_n, \dots les temps de passage successifs d'une chaîne de Markov X par un état fixé $e \in E$. Montrer (utiliser que $\{T_n < \infty\} = \{X_{T_n} = e\}$) que $\forall n \in \mathbb{N}^* : \mathbb{P}_\mu(T_{n+1} - T_n < \infty | T_1 < \infty, T_2 - T_1 < \infty, \dots, T_n - T_{n-1} < \infty) = \mathbb{P}_e(T_1 < \infty)$.

Exercice n° 3.1.8 Étude de la transmission du nom "Chenin", porté à la génération 0 par un unique homme (= humain mâle). Notons Z_n le nombre d'hommes s'appelant "Chenin" à la génération $n \in \mathbb{N}$, F la fonction génératrice de Z_1 ($F(s) := \mathbb{E}(s^{Z_1})$), p_k la probabilité (supposée fixe) qu'un homme ait k enfants mâles, et posons $m := \sum_{\{k \in \mathbb{N}\}} k p_k$. Supposons l'indépendance entre les descendance des différents hommes.

- Calculer par récurrence la fonction G_n , génératrice de Z_n , en fonction de F .
- Vérifier que F est monotone et convexe sur $[0, 1]$.
- En déduire en fonction de m le comportement asymptotique de $\alpha_n := \mathbb{P}(Z_n = 0)$. Interprétation ? On montrera que pour $m \leq 1$ le nom s'éteint presque sûrement.

Exercice n° 3.1.9 Fixons $n \in \mathbb{N}^*$, et un espace d'états $E := \{1, \dots, n\}$, muni d'une matrice de transition P associée à une chaîne de Markov X sur E . Fixons une permutation $\sigma \in \mathcal{S}_n$, et notons $E' := \{\sigma(1), \dots, \sigma(n)\}$ l'espace d'états égal à E en tant qu'ensemble, mais dont les éléments sont écrits dans un autre ordre ; de sorte que la matrice de transition P' associée à la même chaîne de Markov X vue comme chaîne sur E' diffère de P . À la permutation σ associons la matrice $M^\sigma \in \mathcal{M}_n(\mathbb{Z})$ définie par $M_{jk}^\sigma := 1_{\{j=\sigma(k)\}}$. Exprimer la matrice P' en fonction des matrices P et M^σ .

3.2 Classification des états

Définition 9 *Considérons une chaîne de Markov X de matrice de transition P , et deux états $x, y \in E$.*

On dit que x conduit à y ($x \rightarrow y$) lorsqu'il existe $n \in \mathbb{N}$ tel que $P^n(x, y) > 0$.

On dit que x et y communiquent ($x \leftrightarrow y$) lorsque x conduit à y et y conduit à x .

On appelle classe de communication toute classe d'équivalence de E pour la relation \leftrightarrow .

On dit que la chaîne X (ou la matrice P) est irréductible lorsqu'il n'y a qu'une seule classe de communication.

Exercice n° 3.2.1 a) Vérifier que la relation ($x \leftrightarrow y$) est bien une relation d'équivalence sur E . b) Montrer que $\mathbb{P}_x(T_y < \infty) > 0 \iff x \rightarrow y$.

c) Vérifier si les exemples 2) et 3) de la section précédente sont des exemples de chaîne irréductible. Déterminer les classes de communication de la chaîne définie par

$$E := \{1, \dots, 6\} \text{ et } P(1, 2) = P(2, 3) = P(3, 4) = P(4, 5) = P(5, 3) = P(6, 6) = 1.$$

d) Montrer que si une chaîne est irréductible son graphe de transition est connexe, mais que la réciproque est en général fautive.

Définition 10 *On dit qu'une classe de communication C conduit à une classe de communication C' ($C \rightarrow C'$) lorsqu'il existe $x \in C$ et $x' \in C'$ tels que x conduit à x' .*

Une partie A de E est dite fermée lorsque ($x \in A$ et $x \rightarrow y$) $\implies y \in A$, et elle est dite absorbante lorsqu'elle est fermée et lorsque de plus on atteint p.s. A en partant de n'importe quel point de E .

Exercice n° 3.2.2 a) Vérifier que la relation ($C \rightarrow C'$) sur l'ensemble des classes de communication est une relation d'ordre (partiel).

b) Trouver les parties fermées et les parties absorbantes pour l'exemple de l'exercice 3.2.1.c, puis pour la chaîne définie par $E := \{1, \dots, 6\}$ et

$$P(1, 2) = 2P(2, 3) = P(3, 3) = 2P(2, 4) = P(4, 5) = P(5, 6) = P(6, 5) = 1.$$

Définition 11 *On appelle période d'un état fixé $e \in E$:*

$$d_e := \text{p.g.c.d.}(\{n \in \mathbb{N}^* \mid P^n(e, e) > 0\}), \text{ qui vaut } +\infty \text{ si } (\forall n \in \mathbb{N}^*) P^n(e, e) = 0.$$

L'état e est dit apériodique lorsque $d_e = 1$.

Exercice n° 3.2.3 Déterminer les périodes pour les deux exemples de l'exercice 3.2.2.b et pour les marches simples sur \mathbb{Z}^d , ainsi que pour les urnes d'Ehrenfest.

Proposition 8 *Deux états d'une même classe de communication C ont nécessairement la même période, qu'on appelle période de la classe de communication C .*

Preuve Soient x, y deux états distincts qui communiquent. Notons que d_x et d_y sont nécessairement finis, et fixons $n, m, k \in \mathbb{N}^*$ tels que $P^n(x, y) P^m(y, x) P^k(y, y) > 0$.

Alors pour tout $j \in \mathbb{N}$ on a : $P^{m+jk+n}(x, x) \geq P^n(x, y) (P^k(y, y))^j P^m(y, x) > 0$, et donc d_x divise k . Par conséquent d_x divise d_y . \diamond

Lemme 5 Soit x un état de période finie d_x . Il existe un entier n_x tel que $P^{nd_x}(x, x) > 0$ pour tout entier $n > n_x$.

Preuve Fixons $n_1, \dots, n_k \in \mathbb{N}^*$ dont d_x est le p.g.c.d., tels que $P^{n_j}(x, x) > 0$ pour $1 \leq j \leq k$. Il existe un entier n_x tel que tout entier $n > n_x$ s'écrive $n = a_1 \frac{n_1}{d_x} + \dots + a_k \frac{n_k}{d_x}$, avec $a_1, \dots, a_k \in \mathbb{N}$ (cela se déduit de la relation de Bezout, par exemple par récurrence sur k , montrant que pour n, m premiers entre eux et a donnés, tout entier assez grand s'écrit $un + vm$ avec u, v entiers tels que $u, v > a$). Alors $P^{nd_x}(x, x) \geq (P^{n_1}(x, x))^{a_1} \dots (P^{n_k}(x, x))^{a_k} > 0$. \diamond

Définition 12 Un état $e \in E$ est dit récurrent lorsque le temps de retour en e est \mathbb{P}_e -p.s. fini. Il est dit transitoire (ou transient) sinon. Une classe de communication est dite récurrente lorsque tous ses états sont récurrents, et transitoire (ou transiente) sinon.

Exercice n° 3.2.4 Notons $T_e := \inf\{n \in \mathbb{N}^* \mid X_n = e\}$ le temps de retour en $e \in E$, et $N_e := \sum_{\{n \in \mathbb{N}\}} 1_{\{X_n = e\}}$ le nombre total de passages en e .

- Montrer que $\mathbb{P}_e(N_e > k) = \mathbb{P}_e(T_e^k < \infty) = (\mathbb{P}_e(T_e < \infty))^k$, pour tout $k \in \mathbb{N}$.
- Montrer que $\mathbb{E}_e(N_e) = \sum_{\{n \in \mathbb{N}\}} P^n(e, e)$.
- Montrer que si $\mathbb{P}_e(T_e < \infty) = 1$, alors $\mathbb{P}_e(N_e = \infty) = 1$ et $\mathbb{E}_e(N_e) = \infty$.
- Montrer que si $\mathbb{P}_e(T_e < \infty) < 1$, alors e est transitoire et $\mathbb{E}_e(N_e) < \infty$.
- Déduire que les états d'une même classe de communication sont soit tous récurrents, soit tous transitoires. (Indication : $\mathbb{E}_e(N_e) \geq P^k(e, x) \mathbb{E}_x(N_x) P^l(x, e)$.)
- Montrer que toute classe de communication récurrente est fermée. (Indication : si e est récurrent, alors $1 = \sum_y P^n(e, y) \mathbb{P}_y(T_e < \infty)$ pour tout $n \in \mathbb{N}$.)
- Montrer que toute classe de communication fermée et finie est récurrente. En particulier une chaîne irréductible sur E fini est récurrente. (Indication : $\mathbb{E}_x(N_x) \geq P^k(x, e) \mathbb{E}_e(N_e)$.)
- Montrer que la chaîne définie par $E = \mathbb{N}^*$ et $(\forall n \in E) P(n, 1) = 1 - P(n, n+1) = 2^{-n}$ est irréductible et transitoire. (Évaluer $\mathbb{P}_2(T_1 = \infty)$). Quel autre exemple a déjà été vu ?

Remarquons que a,b,c,d de l'exercice 3.2.4 ci-dessus entraînent aussitôt :

Proposition 9 Fixons un état $e \in E$, puis $T_e := \inf\{n \in \mathbb{N}^* \mid X_n = e\}$ et $N_e := \sum_{\{n \in \mathbb{N}\}} 1_{\{X_n = e\}}$. Nous avons :

$$e \text{ récurrent} \Leftrightarrow \mathbb{P}_e(T_e < \infty) = 1 \Leftrightarrow \mathbb{P}_e(N_e = \infty) = 1 \Leftrightarrow \mathbb{E}_e(N_e) = \sum_n P^n(e, e) = \infty, \quad \text{et}$$

$$e \text{ transitoire} \Leftrightarrow \mathbb{P}_e(T_e < \infty) < 1 \Leftrightarrow \mathbb{P}_e(N_e < \infty) = 1 \Leftrightarrow \mathbb{E}_e(N_e) = \sum_n P^n(e, e) < \infty.$$

Réénonçons le résultat de l'exercice 3.2.4.e ci-dessus.

Proposition 10 Les états d'une même classe de communication sont soit tous récurrents, soit tous transitoires.

Proposition 11 *Supposons la chaîne irréductible et récurrente. Alors tous les temps d'atteinte ou de retour sont p.s. finis.*

Preuve Fixons $x \in E$. Puisque $\mathbb{P}_\mu(T_x < \infty) = \sum_{\{y \in E\}} \mu(y) \mathbb{P}_y(T_x < \infty)$, il suffit de montrer que $\mathbb{P}_y(T_x < \infty) = 1$ pour tout $y \in E$. Si $y = x$, c'est dans la proposition 9. Supposons donc $y \neq x$, et fixons $n \in \mathbb{N}^*$ tel que $P^n(x, y) > 0$.

Utilisant la proposition 9, nous avons :

$$\begin{aligned} \sum_{e \in E} P^n(x, e) &= 1 = \mathbb{P}_x(N_x = \infty) = \mathbb{P}_x((\exists m > n) X_m = x) \\ &= \sum_{e \in E} \mathbb{P}_x((\exists m > n) X_m = x \mid X_n = e) \mathbb{P}_x(X_n = e) = \sum_{e \in E} \mathbb{P}_e(T_x < \infty) P^n(x, e), \end{aligned}$$

d'où $0 = \sum_e (1 - \mathbb{P}_e(T_x < \infty)) P^n(x, e) \geq (1 - \mathbb{P}_y(T_x < \infty)) P^n(x, y)$. \diamond

Réénonçons le résultat de l'exercice 3.2.4.g ci-dessus.

Proposition 12 *Si E est fini, il contient au moins un point récurrent.*

Preuve Supposons le contraire. Alors d'après la proposition 9, nous avons pour tout $x \in E$: $\mathbb{P}_x(N_x = \infty) = 0$. Soient $x, y \in E$. Utilisant la proposition 7, nous avons :

$$\mathbb{P}_y(N_x = \infty) = \mathbb{P}_y(T_x < \infty, N_x \circ \theta^{T_x} = \infty) = \mathbb{P}_x(N_x = \infty) \mathbb{P}_y(T_x < \infty) = 0,$$

d'où $\mathbb{P}_y((\exists x \in E) N_x = \infty) = 0$, et donc $\mathbb{P}_\mu((\exists x \in E) N_x = \infty) = 0$, d'où enfin $1 = \mathbb{P}_\mu((\forall x \in E) N_x < \infty) = \mathbb{P}_\mu(\sum_{\{x \in E\}} N_x < \infty) = \mathbb{P}_\mu(\infty < \infty) = 0$. \diamond

Exercice n° 3.2.5 Supposons E de cardinal fini $n \in \mathbb{N}^*$. $T_y := \min\{k \in \mathbb{N}^* \mid X_k = y\}$.

a) Montrer que si $x \rightarrow y$, alors $\mathbb{P}_x(T'_y < n) > 0$, où $T'_y := \min\{k \in \mathbb{N} \mid X_k = y\}$.

b) Supposons E irréductible. Dédurre de a) qu'il existe $\alpha \in]0, 1[$ tel que

$\mathbb{P}_x(T_y \geq k(n+1)) \leq \alpha^k$ pour tous $k \in \mathbb{N}$ et $x, y \in E$; puis qu'il existe $\beta \in]0, 1[$ et $c > 0$ tels que $\mathbb{P}_x(T_y \geq m) \leq c \beta^m$ pour tous $m > n$ et $x, y \in E$.

c) Dédurre que tout temps d'atteinte ou de retour possède un moment exponentiel.

Exercice n° 3.2.6 Soit $\{b_n \mid n \in \mathbb{N}\}$ une suite décroissante de réels > 0 telle que $b_0 = 1$, et soient $\sigma_n := \sum_{j=0}^n b_j$ et $\beta_n := b_n / \sigma_n$, pour tout $n \in \mathbb{N}$. Considérons la matrice de transition P sur $E = \mathbb{N}$ définie par $P(i, j) := \frac{b_i}{b_i} (\beta_i - \beta_{i+1}) 1_{\{j \leq i\}} + \frac{\beta_{i+1}}{\beta_i} 1_{\{j=i+1\}}$, pour tous i, j dans \mathbb{N} . Montrer que P est irréductible, et que $P^n(0, i) = \frac{b_i}{\sigma_n} 1_{\{i \leq n\}}$ pour tous i, n dans \mathbb{N} . Dédurre quand la chaîne correspondante est récurrente.

3.3 Mesures invariantes des chaînes finies

Définition 13 Une mesure ν sur E est dite *invariante* (ou *stationnaire*) pour la chaîne définie par la matrice P lorsque $\nu P = \nu$.

Remarquons que si la loi initiale ν de X est invariante pour P , alors toutes les variables X_n ont aussi ν pour loi, et pour tout $n \in \mathbb{N}$ la chaîne $X \circ \theta^n$ a la même loi que X .

Exercice n° 3.3.1 Supposons que E est fini et que, pour un certain $e \in E$, $P^n(e, x)$ converge lorsque $n \rightarrow \infty$ vers un certain p_x , pour chaque $x \in E$. Montrer que $(p_x | x \in E)$ est une probabilité invariante pour P . Montrer que la finitude de E est ici une hypothèse indispensable.

Exercice n° 3.3.2 Chercher les mesures invariantes des chaînes suivantes :

- a) $E = \mathbb{N}$ et $P(n, n+1) = 1$, pour tout $n \in \mathbb{N}$.
- b) $E = \mathbb{N}$ et $P(n, 0) = p_n = 1 - P(n, n+1)$, pour tout $n \in \mathbb{N}$. Discuter suivant les p_n .
- c) $E = \mathbb{Z}$ et $P(n, n-1) = p = 1 - P(n, n+1)$, pour tout $n \in \mathbb{Z}$. Discuter suivant p .

Proposition 13 Les points transitoires de E fini sont négligeables pour toute mesure invariante. C'est faux en général si E est infini.

Preuve Soit ν une probabilité invariante chargeant un point transitoire. Notons \mathcal{T} l'ensemble des points transitoires chargés par ν . Supposons que $\sum_{\{t \in \mathcal{T}\}} P(e, t) = 1$ pour tout $e \in \mathcal{T}$; alors la chaîne issue d'un point de \mathcal{T} resterait dans \mathcal{T} , qui est fini, et la proposition 12 imposerait l'existence d'un point récurrent dans \mathcal{T} , ce qui est contradictoire. Donc il existe $e \in \mathcal{T}$ tel que $\sum_{\{t \in \mathcal{T}\}} P(e, t) < 1$. Cela entraîne finalement (avec 3.2.4.f) :

$$\begin{aligned} \nu(\mathcal{T}) &= \sum_{t \in \mathcal{T}} \nu(t) = \sum_{t \in \mathcal{T}} \sum_{e \in E} \nu(e) P(e, t) = \sum_{e \in E} \nu(e) \sum_{t \in \mathcal{T}} P(e, t) = \sum_{e \in \mathcal{T}} \nu(e) \sum_{t \in \mathcal{T}} P(e, t) \\ &< \sum_{\{e \in \mathcal{T}\}} \nu(e) = \nu(\mathcal{T}) . \diamond \end{aligned}$$

Proposition 14 Sur E fini, il y a au moins une probabilité invariante pour P .

Preuve Fixons une probabilité quelconque μ sur E , et posons $\mu_n := \frac{1}{n} \sum_{k=0}^{n-1} \mu P^k$, pour tout $n \in \mathbb{N}^*$. C'est une suite de probabilités sur E . Or l'ensemble des probabilités sur E s'identifie à un fermé dans $[0, 1]^{\text{Card}(E)}$, et donc est compact. La suite μ_n admet donc une sous-suite μ_{n_k} convergente, vers une certaine probabilité ν .

Par ailleurs nous avons pour tout $k \in \mathbb{N}^*$: $\mu_{n_k} P = \mu_{n_k} + (\mu P^{n_k} - \mu) / n_k$, d'où $\nu P = \nu$ en passant à la limite. \diamond

Théorème 5 Sur E fini, il y a une seule probabilité invariante pour P si et seulement si il n'y a qu'une seule classe de communication récurrente.

Preuve S'il y a au moins deux classes récurrentes C et C' , la restriction de la chaîne à C admet une probabilité invariante ν , d'après l'énoncé précédent. De même, nous avons ν' pour C' . De sorte que les deux probabilités ν et ν' sont distinctes et invariantes pour P . Pour la réciproque, la proposition 13 et l'exercice 3.2.4.f permettent de se restreindre au cas où P est récurrente irréductible. Il suffit d'établir que le sous-espace propre de la matrice tP associé à la valeur propre 1 est une droite. Or on voit par réduction de Jordan (ou bien : comme $\text{Ker}\Delta = (\text{Im } {}^t\Delta)^\perp$, $\dim(\text{Ker}\Delta) = \dim E - \dim(\text{Im } {}^t\Delta) = \dim[\text{Ker}({}^t\Delta)]$) que les noyaux de $P - \mathbf{1}$ et ${}^tP - \mathbf{1}$ ont la même dimension. Il suffit donc d'établir que le sous-espace propre de la matrice P associé à la valeur propre 1 est une droite ; c'est-à-dire que toute fonction f sur E telle que $Pf = f$ est nécessairement constante. Fixons une telle fonction, et posons pour tout $x \in E$: $F_n(x) := \sum_{y \in E} P^n(x, y)(f(x) - f(y))^2$, pour tout $n \in \mathbb{N}$.

Développant le carré, on voit aussitôt que $F_n = f^2 - 2f P^n f + P^n(f^2) = P^n(f^2) - f^2$. Fixons maintenant une probabilité invariante ν .

Observons que $\nu(x) \geq \nu(y)P^n(y, x)$ pour tous $x, y \in E$, et donc que $[\nu(x) = 0 \text{ et } P^n(y, x) > 0] \Rightarrow \nu(y) = 0$. Par conséquent $[\nu(x) = 0 \text{ et } y \rightarrow x] \Rightarrow \nu(y) = 0$, ce qui prouve que ν charge tous les points de E . Finalement pour tout $n \in \mathbb{N}$

$$0 = \int (P^n(f^2) - f^2) d\nu = \int F_n d\nu = \sum_{x, y \in E} \nu(x) P^n(x, y)(f(x) - f(y))^2$$

montre que $P^n(x, y) > 0 \Rightarrow f(x) = f(y)$, d'où $x, y \in E \Rightarrow x \rightarrow y \Rightarrow f(x) = f(y)$. \diamond

Corollaire 4 *S'il n'y a qu'une classe récurrente, pour toute probabilité μ sur E fini, la suite $\mu_n := (\mu + \mu P + \dots + \mu P^{n-1})/n$ converge vers la probabilité invariante.*

Exercice n° 3.3.3 Dédurre des résultats ci-dessus et de leur preuves que s'il y a exactement k classes de communication récurrentes, alors le cône des mesures positives invariantes est de dimension k . Dédurre aussi que si l'on suppose de plus qu'il n'y a pas de point transitoire, alors l'espace des fonctions invariantes (i.e. f telles que $Pf = f$) est aussi de dimension k .

Exercice n° 3.3.4 Soit P une matrice stochastique sur $E = \{1, \dots, n\}$. Montrer qu'une loi ν sur E est invariante ssi $\nu(1 - P + U) = u$, où 1 désigne la matrice unité, U désigne la matrice ne comportant que des 1, et u désigne le vecteur-ligne ne comportant que des 1. Dédurre que si P est irréductible, alors la matrice $(1 - P + U)$ est inversible (et donc $\nu = u(1 - P + U)^{-1}$). (Indication : considérer une loi invariante ν , et ${}^t\mu \in \text{ker } {}^t(1 - P + U)$.)

Exercice n° 3.3.5 Soit P la matrice stochastique sur $E = \{1, \dots, 5\}$ donnée par

$$P := \begin{pmatrix} 1/3 & 1/2 & 0 & 1/6 & 0 \\ 1/4 & 0 & 0 & 3/4 & 0 \\ 1/8 & 0 & 1/4 & 1/8 & 1/2 \\ 1/4 & 0 & 0 & 3/4 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \end{pmatrix}.$$

Préciser pour la chaîne de Markov définie par P sur E : les classes de communication, les classes récurrentes, les périodes, les mesures invariantes.

3.4 Probabilité invariante

Revenons au cas général : nous n'imposons plus ici que E soit fini.

Définition 14 *Un état récurrent $e \in E$ est dit récurrent nul lorsque $\mathbb{E}_e(T_e)$ est infini, et récurrent positif sinon. Il est dit ergodique lorsqu'il est récurrent positif et apériodique.*

Proposition 15 *Soit $e \in E$ un état d'une chaîne irréductible récurrente de matrice P . Considérons le vecteur-ligne (donc la mesure sur E) V^e dont le terme d'indice $x \in E$ est $V_x^e := \mathbb{E}_e\left(\sum_{n=1}^{T_e} 1_{\{X_n=x\}}\right) = \sum_{n=1}^{\infty} \mathbb{P}_e(X_n = x, T_e \geq n)$. Alors tous les V_x^e sont finis et > 0 , $V_e^e = 1$, $\mathbb{E}_e(T_e) = V^e(E)$, et V^e est une mesure invariante ($V^e = V^e P$).*

Preuve Observons d'abord que $V_e^e = 1$, et que $T_e = \sum_{x \in E} \sum_{n=1}^{T_e} 1_{\{X_n=x\}}$, de sorte que $\mathbb{E}_e(T_e) = \sum_{x \in E} V_x^e = V^e(E)$. Ensuite, pour $x \neq e$ et $n \in \mathbb{N}^*$, fixant $m \in \mathbb{N}^*$ tel que $\mathbb{P}_x(T_e = m) > 0$ (son existence découle de la proposition 11), nous avons :

$\mathbb{P}_e(T_e = m + n) = \mathbb{P}_e(T_e \geq n, X_n \neq e, T_e \circ \theta^n = m) \geq \mathbb{P}_e(T_e \geq n, X_n = x) \mathbb{P}_x(T_e = m)$, de sorte que

$$V_x^e \leq \sum_{n=1}^{\infty} \mathbb{P}_e(T_e = m + n) / \mathbb{P}_x(T_e = m) \leq \mathbb{P}_x(T_e = m)^{-1} < \infty .$$

L'argument déjà employé dans la preuve du théorème 5 montre que l'irréductibilité de P , l'invariance de V^e et le fait que $V_e^e > 0$ entraînent que $V_x^e > 0$ pour tout $x \in E$.

Pour la dernière assertion, observons que $\mathbb{P}_e(X_1 = x, T_e \geq 1) = P(e, x)$, et que $\mathbb{P}_e(X_n = x, T_e \geq n) = \sum_{y \neq e} \mathbb{P}_e(X_{n-1} = y, T_e \geq n-1) P(y, x)$, pour $n \geq 2$, ce qui donne par sommation :

$$\begin{aligned} V_x^e &= P(e, x) + \sum_{y \neq e} \sum_{n \geq 2} \mathbb{P}_e(X_{n-1} = y, T_e \geq n-1) P(y, x) = V_e^e P(e, x) + \sum_{y \neq e} V_y^e P(y, x) \\ &= (V^e P)_x . \quad \diamond \end{aligned}$$

Corollaire 5 *Si la chaîne est irréductible récurrente sur E , elle admet une mesure invariante dont le support est E entier. Si de plus E contient un état récurrent positif, alors il existe une probabilité invariante dont le support est E entier.*

Exercice n° 3.4.1 Soit X une chaîne irréductible récurrente de matrice P sur E . Fixons $e \in E$. Soit ν une mesure non nulle invariante pour P , normalisée pour que $\nu_e = 1$. Pour tous x, y dans E , posons $Q(x, y) := P(y, x) \times \nu_y / \nu_x$.

a) Montrer que Q est une matrice stochastique, calculer $Q^n(x, y)$ pour tout $n \in \mathbb{N}^*$, et montrer que Q est irréductible récurrente. Comparer les périodes de (E, P) et de (E, Q) .

- b) Caractériser les mesures P -invariantes au moyen de leur densité par rapport à ν et de Q . Montrer que (E, Q) est nulle-récurrente ssi (E, P) l'est.
- c) Dédire une variante de la preuve du théorème 5.
- d) Vérifier que $Q^n(x, y) = \mathbb{P}_\nu(X_{N-n} = y | X_N = x)$ pour tous n, N entiers tels que $0 \leq n \leq N$ et tous x, y dans E : Q est associée à la chaîne X "retournée".
- e) Montrer par récurrence sur $n \in \mathbb{N}^*$ que pour tout $x \in E \setminus \{e\}$ on a $\nu_x \mathbb{Q}_x(T_e = n) = \mathbb{P}_e(X_n = x, T_e > n)$, où \mathbb{Q} désigne la loi de la chaîne associée à Q .

Théorème 6 *Une chaîne irréductible admet une probabilité invariante si et seulement si tous ses états sont récurrents positifs, ou bien si et seulement si l'un de ses états est récurrent positif. Dans ce cas, la probabilité invariante est unique, et confère le poids $\mathbb{E}_e(T_e)^{-1}$ à chaque état $e \in E$.*

Preuve Le corollaire 5 ci-dessus assure l'existence d'une probabilité invariante s'il y a un état récurrent positif. Réciproquement, soit ν une probabilité invariante. Si les états sont transitoires, alors la proposition 9 (et l'exercice 3.2.4.b,g) montrent que tous les $P^n(y, x)$ tendent vers 0 lorsque $n \rightarrow \infty$; et donc pour tout $x \in E$: $\nu_x = \sum_y \nu_y P^n(y, x) \rightarrow 0$ par convergence dominée, ce qui est contradictoire. La chaîne est donc récurrente.

Plaçons-nous ensuite dans le cas où la loi de X_0 est ν , et fixons $e \in E$.

Rappelons d'abord un argument déjà utilisé : si jamais nous avions $\nu_e = 0$, alors $0 = \nu_e = \sum_x \nu_x P^n(x, e)$ pour tout $n \in \mathbb{N}$ montre que nous aurions $\nu_x = 0$ pour tout $x \in E$ tel que $x \rightarrow e$, et donc $\nu = 0$, ce qui est exclu ; par conséquent $\nu_e > 0$.

Cela dit, nous avons :

$$\begin{aligned}
\nu_e \mathbb{E}_e(T_e) &= \sum_{n \in \mathbb{N}} \mathbb{P}_\nu(X_0 = e) \mathbb{P}_\nu(T_e > n | X_0 = e) = \sum_{n \in \mathbb{N}} \mathbb{P}_\nu(T_e > n, X_0 = e) \\
&= \mathbb{P}_\nu(X_0 = e) + \sum_{n \in \mathbb{N}^*} \mathbb{P}_\nu(X_0 = e, X_1 \neq e, \dots, X_n \neq e) \\
&= \mathbb{P}_\nu(X_0 = e) + \sum_{n \in \mathbb{N}^*} \left[\mathbb{P}_\nu(X_1 \neq e, \dots, X_n \neq e) - \mathbb{P}_\nu(X_0 \neq e, \dots, X_n \neq e) \right] \\
&= \mathbb{P}_\nu(X_0 = e) + \sum_{n \in \mathbb{N}^*} \left[\mathbb{P}_\nu(X_0 \neq e, \dots, X_{n-1} \neq e) - \mathbb{P}_\nu(X_0 \neq e, \dots, X_n \neq e) \right] \\
&= \mathbb{P}_\nu(X_0 = e) + \mathbb{P}_\nu(X_0 \neq e) - \mathbb{P}_\nu(\forall n \in \mathbb{N}) X_n \neq e = 1,
\end{aligned}$$

par la proposition 11. Ceci prouve à la fois que l'état (quelconque) e est récurrent positif, puisque $\mathbb{E}_e(T_e) = (\nu_e)^{-1} < \infty$, et que la probabilité invariante (quelconque) ν est bien nécessairement égale à ce qu'énonce le théorème. \diamond

Définition 15 *Une chaîne irréductible récurrente est dite positivement récurrente lorsque tous ses états sont récurrents positifs, et sinon, lorsqu'aucun ne l'est, elle est dite nulle-récurrente.*

Proposition 16 *Supposons la chaîne irréductible positivement récurrente. Alors (pour tout $e \in E$) tous les temps d'atteinte sont \mathbb{P}_e -intégrables.*

Preuve Fixons $x \neq y$ dans E . Nous savons déjà que tous les temps d'atteinte sont p.s. finis (proposition 11), et que les temps de retour sont intégrables.

Supposons qu'on ait $T_x > T_y^k$ \mathbb{P}_y -p.s. , pour un $k \in \mathbb{N}^*$; on aurait alors aussi par la propriété de Markov forte $T_x \circ \theta^{T_y^k} > T_y \circ \theta^{T_y^k}$ \mathbb{P}_y -p.s. , et donc \mathbb{P}_y -p.s. :

$$T_x = T_y^k + T_x \circ \theta^{T_y^k} > T_y^k + T_y \circ \theta^{T_y^k} = T_y^{k+1}, \text{ d'où par récurrence, si } T_x > T_y \text{ } \mathbb{P}_y\text{-p.s. :}$$

$$T_x > T_y^k \geq k \text{ pour tout } k \in \mathbb{N}^*, \text{ et donc } T_x = \infty. \text{ Donc il faut que } \mathbb{P}_y(T_x < T_y) > 0.$$

Ensuite $T_y \geq (T_x + T_y \circ \theta^{T_x}) 1_{\{T_x < T_y\}} \geq T_y \circ \theta^{T_x} 1_{\{T_x < T_y\}}$, et donc par la propriété de Markov forte $\mathbb{E}_y(T_y) \geq \mathbb{P}_y(T_x < T_y) \times \mathbb{E}_x(T_y)$, d'où $\mathbb{E}_x(T_y) < \infty$. \diamond

Exercice n° 3.4.2 Pour les exemples 2) (urnes d'Ehrenfest) et 3) de la section 3.1, chercher les mesures invariantes, et conclure quant à la récurrence (positive).

Pour l'exemple 3) ($E = \mathbb{N}$, $P(n, n-1) = 1$ si $n \geq 1$, $P(0, n) = \mathbb{P}(Y = n) > 0$ si $n \geq 0$), on vérifiera que la récurrence est positive ssi $Y \in L^1$ et qu'alors, ν étant invariante, $\mathbb{E}_\nu(T'_0)$ est fini ssi $Y \in L^2$. En particulier, dans la proposition 16 ci-dessus, on ne peut pas avoir l'intégrabilité des temps d'atteinte sous n'importe quelle loi initiale.

De même, chercher les mesures invariantes de la marche sur \mathbb{Z} définie par $P(j, j \pm 1) = P(j, j \pm 2) = 1/4$; vérifier que c'est une chaîne irréductible, apériodique, nulle-récurrente.

Exercice n° 3.4.3 Une particule se déplace sur les huit sommets d'un cube, choisissant à chaque seconde (avec loi uniforme) l'un des trois sommets voisins. Fixons deux sommets opposés du cube, disons e et a , et faisons partir la particule de e au temps 0.

- Quel est le temps moyen de retour en e ?
- Quel est le nombre moyen de passages en a avant le retour en e ?
- Notons B l'ensemble des trois sommets voisins de e , et A l'ensemble des trois sommets voisins de a . Si la particule est en X_n au temps n , posons $Y_n := B 1_B(X_n) + A 1_A(X_n) + X_n 1_{\{X_n=e \text{ ou } a\}}$. Vérifier que Y est une chaîne de Markov sur $\{e, a, A, B\}$, et donner sa matrice de transition. (On dit qu'on a pu "agrégé" les états de A et de B ; ce qui n'est en général pas possible).
- Relativement à la chaîne Y , calculer $\mathbb{E}_a(T_a)$, puis $\mathbb{E}_e(T_a)$, $\mathbb{E}_B(T_a)$ et $\mathbb{E}_A(T_a)$.
- Quel est le temps moyen nécessaire à la particule pour atteindre a ?

Exercice n° 3.4.4 Supposons que l'état $e \in E$ est récurrent positif pour la chaîne X issue de e , et notons G la fonction génératrice de T_e : $G(t) := \mathbb{E}_e(t^{T_e})$, pour $|t| < 1$.

Soit $Y_n := \max\{k \leq n \mid X_k = e\}$ le dernier instant de passage en e avant l'instant n .

- Montrer que $\sum_{k \in \mathbb{N}} t^k \mathbb{P}_e(T_e > k) = (1 - G(t))/(1 - t)$, pour $|t| < 1$.
- Montrer que $\mathbb{P}_e(Y_n = j) = \sum_{k \in \mathbb{N}} \mathbb{P}_e(T_e^k = j) \mathbb{P}_e(T_e > n - j)$, pour $0 \leq j \leq n$.
- Déduire que $\sum_{n \in \mathbb{N}} t^n \mathbb{E}_e(s^{Y_n}) = \frac{1 - G(t)}{(1-t)(1-G(st))}$, pour $|s|, |t| < 1$. (Changer n en $n + j$.)

3.5 Ergodicité

Exercice n° 3.5.1 Montrer que pour tous états $x \neq y$ on a :

$$\sum_{n \in \mathbb{N}} P^n(x, y) = \mathbb{E}_x(N_y) = \mathbb{P}_x(T_y < \infty) \mathbb{E}_y(N_y). \quad (\text{Si } x = y, \text{ voir l'exercice 3.2.4.b}).$$

Lemme 6 Si ν est une probabilité invariante chargeant l'état y , cet état est récurrent.

Preuve Utilisant l'exercice 3.5.1 ci-dessus, nous avons : $\infty = \sum_n \nu_y =$

$$\begin{aligned} \sum_n \sum_x \nu_x P^n(x, y) &= \sum_x \nu_x \sum_n P^n(x, y) = \sum_x \nu_x \mathbb{P}_x(T_y < \infty) \mathbb{E}_y(N_y) = \mathbb{P}_\nu(T_y < \infty) \mathbb{E}_y(N_y) \\ &\leq \mathbb{E}_y(N_y), \text{ et nous concluons par la proposition 9. } \diamond \end{aligned}$$

Proposition 17 Considérons deux états $e, x \in E$. Alors

$$\frac{1}{n} \sum_{k=1}^n 1_{\{X_k=e\}} \longrightarrow \mathbb{E}_e(T_e)^{-1} \times 1_{\{T_e < \infty\}} \quad \mathbb{P}_x - \text{p.s.} \quad \text{lorsque } n \rightarrow \infty.$$

Preuve D'abord, si $x = e$, notons $N_e^n := \sum_{k=1}^n 1_{\{X_k=e\}} = \max\{j \in \mathbb{N} \mid T_e^j \leq n\}$, et utilisons la proposition 9. Si e est récurrent, alors \mathbb{P}_e -p.s. :

$N_e^n \rightarrow \infty$; $T_e^{N_e^n} \leq n < T_e^{N_e^n+1}$; et les $T_e^k - T_e^{k-1}$ étant i.i.d. > 0 , la loi des grands nombres assure que T_e^k/k converge vers $\mathbb{E}_e(T_e)$ lorsque $k \rightarrow \infty$; d'où nous déduisons que $\frac{n}{N_e^n}$ converge \mathbb{P}_e -p.s. vers $\mathbb{E}_e(T_e)$. Tandis que si e est transitoire alors d'une part $\mathbb{P}_e(T_e = \infty) > 0$ et donc $\mathbb{E}_e(T_e) = \infty$, et d'autre part p.s. N_e^n est borné et donc $N_e^n/n \rightarrow 0$. Donc si $x = e$ le résultat est valable, avec ou sans l'indicatrice de $\{T_e < \infty\}$.

Lorsque $x \neq e$, le résultat est clair sur $\{T_e = \infty\}$, et pour tout $m \in \mathbb{N}^*$ (tel que $\mathbb{P}_x(T_e = m) > 0$) on a $\mathbb{P}_x(N_e^n/n \rightarrow \mathbb{E}_e(T_e)^{-1} \mid T_e = m) = \mathbb{P}_e(N_e^{n-m}/n \rightarrow \mathbb{E}_e(T_e)^{-1}) = 1$, d'où le résultat également sur $\{T_e < \infty\}$. \diamond

Corollaire 6 Pour tous $x, y \in E$, nous avons lorsque $n \rightarrow \infty$:

$$\frac{1}{n} \sum_{k=1}^n P^k(x, y) \longrightarrow \mathbb{P}_x(T_y < \infty) / \mathbb{E}_y(T_y).$$

Il est en général faux qu'on puisse se passer de prendre la moyenne de Césaro des $P^n(x, y)$ pour avoir le résultat ci-dessus. Par exemple si P est la matrice (2,2) avec 0 sur la diagonale. Pour contrer le phénomène de périodicité qu'on voit ainsi s'opposer au résultat fort dit d'ergodicité, c'est justement l'hypothèse d'apériodicité qui est la bonne.

Théorème 7 Soit X une chaîne irréductible, apériodique (i.e. tout état est apériodique) et admettant une probabilité invariante ν . Alors nous avons $P^n(x, y) \rightarrow \nu(y)$ lorsque $n \rightarrow \infty$, uniformément par rapport à $y \in E$, pour tout $x \in E$.

Preuve Procédons par couplage ; définissons la chaîne produit \bar{X} sur E^2 en posant : $\bar{P}((x, x'), (y, y')) = P(x, y)P(x', y')$, et vérifions que l'irréductibilité et l'apériodicité de X entraînent l'irréductibilité de \bar{X} . Fixons j, k entiers tels que $P^j(x, y)P^k(x', y') > 0$. Le lemme 5 assure que $P^{k+m}(y, y)P^{j+m}(y', y') > 0$, pour m entier suffisamment grand. Donc

$$\begin{aligned} \bar{P}^{j+k+m}((x, x'), (y, y')) &= P^{j+k+m}(x, y) P^{j+k+m}(x', y') \\ &\geq P^j(x, y)P^{k+m}(y, y)P^k(x', y')P^{j+m}(y', y') > 0. \end{aligned}$$

Observons ensuite que $\bar{\nu}_{(x,y)} := \nu_x \nu_y$ définit une probabilité invariante pour \bar{X} , pour laquelle tous les états de E^2 sont récurrents, selon le lemme 6 (rappelons que nécessairement le support de ν est E entier) ou le théorème 6.

Notons $\bar{X}_n = (Y_n, Z_n)$, et σ le temps d'atteinte de la diagonale. Puisque \bar{X} est irréductible et récurrente, tout point de la diagonale est p.s. atteint (voir la proposition 11), et donc a fortiori $\sigma < \infty$ p.s.. Ensuite, \bar{X} est une chaîne de Markov (à coordonnées indépendantes ayant la loi de X), et pour $n \in \mathbb{N}$ et $y \in E$ nous avons sous toute loi initiale μ pour \bar{X}_0 :

$$\begin{aligned} \bar{\mathbb{P}}_\mu(Y_n = y, \sigma \leq n) &= \sum_{m=0}^n \sum_{x \in E} \bar{\mathbb{P}}_\mu(\sigma = m, Y_m = x, Y_n = y) \\ &= \sum_{m=0}^n \sum_{x \in E} \bar{\mathbb{P}}_\mu(\sigma = m, Y_m = Z_m = x) \bar{\mathbb{P}}_{(x,x)}(Y_{n-m} = y) \\ &= \sum_{m=0}^n \sum_{x \in E} \bar{\mathbb{P}}_\mu(\sigma = m, Z_m = Y_m = x) \bar{\mathbb{P}}_{(x,x)}(Z_{n-m} = y) = \bar{\mathbb{P}}_\mu(Z_n = y, \sigma \leq n). \end{aligned}$$

Donc

$$\bar{\mathbb{P}}_\mu(Y_n = y) = \bar{\mathbb{P}}_\mu(Z_n = y, \sigma \leq n) + \bar{\mathbb{P}}_\mu(Y_n = y, \sigma > n) \leq \bar{\mathbb{P}}_\mu(Z_n = y) + \bar{\mathbb{P}}_\mu(Y_n = y, \sigma > n)$$

et $\bar{\mathbb{P}}_\mu(Z_n = y) \leq \bar{\mathbb{P}}_\mu(Y_n = y) + \bar{\mathbb{P}}_\mu(Z_n = y, \sigma > n)$, d'où

$$\left| \bar{\mathbb{P}}_\mu(Y_n = y) - \bar{\mathbb{P}}_\mu(Z_n = y) \right| \leq \bar{\mathbb{P}}_\mu(Y_n = y, \sigma > n) + \bar{\mathbb{P}}_\mu(Z_n = y, \sigma > n),$$

et $\sum_y \left| \bar{\mathbb{P}}_\mu(Y_n = y) - \bar{\mathbb{P}}_\mu(Z_n = y) \right| \leq 2 \bar{\mathbb{P}}_\mu(\sigma > n)$.

Si $Y_0 = x$ et Z_0 a ν pour loi, i.e. $\mu = \delta_x \otimes \nu$, cela donne la conclusion :

$$\sup_y \left| P^n(x, y) - \nu_y \right| \leq \sum_{y \in E} \left| P^n(x, y) - \nu_y \right| \leq 2 \bar{\mathbb{P}}_{\delta_x \otimes \nu}(\sigma > n) \rightarrow 0. \diamond$$

Exercice n° 3.5.2 1) Montrer que les valeurs propres d'une matrice stochastique (sur E fini) sont de module ≤ 1 .

2) Montrer que les valeurs propres $\neq 1$ d'une matrice stochastique (sur E fini) irréductible et apériodique sont de module < 1 , et que la multiplicité de la valeur propre 1 est 1.

3) Réciproquement, si les valeurs propres d'une matrice stochastique P (sur E de cardinal n) sont $(1, \lambda_2, \dots, \lambda_n)$ avec $|\lambda_j| < 1$ pour $2 \leq j \leq n$, alors P^n converge (vers une matrice dont les lignes sont identiques), et (E, P) ne possède qu'une classe de communication récurrente, nécessairement apériodique (et une seule probabilité invariante).

Exercice n° 3.5.3 On lance n fois un dé ordinaire, dont chaque face a une probabilité > 0 de sortir, et on note S_n la somme des résultats. Soit $m \in \mathbb{N}^*$. Trouver la limite quand $n \rightarrow \infty$ de la probabilité que S_n soit multiple de m . Quelle(s) hypothèse(s) plus faible(s) suffirai(en)t à garantir le même résultat ?

Exercice n° 3.5.4 Soit μ une loi sur \mathbb{N}^* , telle que le p.g.c.d. de son support est 1. Soit $\{X_n | n \in \mathbb{N}^*\}$ une suite de v.a.i.i.d. intégrables de loi μ . Pour tout $n \in \mathbb{N}$, posons

$$\sigma_n := \min\{k \in \mathbb{N} | X_1 + \dots + X_k \geq n\}, \text{ et } Y_n := X_1 + \dots + X_{\sigma_n} - n.$$

a) Montrer que Y est une chaîne de Markov sur \mathbb{N} , et donner sa matrice de transition.

(Distinguer suivant que $Y_n = 0$ ou $Y_n > 0$.)

b) Calculer la limite lorsque $n \rightarrow \infty$ de $\mathbb{P}(Y_n = 0)$. (On pourra revenir à l'exercice 3.4.2.)

c) Dédurre le théorème du renouvellement :

$$\lim_{n \rightarrow \infty} \mathbb{P}((\exists k \in \mathbb{N}^*) X_1 + \dots + X_k = n) = \mathbb{E}(X_1)^{-1}.$$

d) Que se passe-t-il si on rajoute la possibilité que $1 > \mu(\{0\}) > 0$? (et si $X_1 \notin L^1$?)

4 Martingales à temps discret

4.1 Surmartingales, sousmartingales, théorème d'arrêt

Donnons-nous un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une filtration, i.e. d'une suite croissante $\{\mathcal{F}_n | n \in \mathbb{N}\}$ de sous-tribus de \mathcal{F} .

Définition 16 Une suite $\{X_n | n \in \mathbb{N}\}$ de v.a.r. intégrables est une surmartingale lorsque (i) elle est adaptée, i.e. X_n est \mathcal{F}_n -mesurable, pour tout $n \in \mathbb{N}$;

(ii) $\mathbb{E}(X_{n+1} | \mathcal{F}_n) \leq X_n$ p.s. pour tout $n \in \mathbb{N}$.

La suite (X_n) est une sousmartingale lorsque $(-X_n)$ est une surmartingale, et est une martingale lorsqu'elle est à la fois une surmartingale et une sousmartingale.

Remarque : Si (X_n) est une surmartingale, la suite de ses espérances est décroissante !

Exemples : 1) Si X est une v.a.r. intégrable, $X_n := \mathbb{E}(X | \mathcal{F}_n)$ définit une martingale.

2) Soit $\{Y_n | n \in \mathbb{N}\}$ une suite de v.a.r. intégrables indépendantes, et soient $X_n := \sum_{j=0}^n Y_j$, puis $\mathcal{F}_0 := \{\emptyset, \Omega\}$, et $\mathcal{F}_n := \sigma\{Y_j | 1 \leq j \leq n\}$. Alors (X_n) est une martingale ssi les

Y_j sont centrées (et donc en particulier les marches aléatoires centrées sont à la fois des martingales et des chaînes de Markov), (X_n) est une surmartingale ssi les Y_j ont toutes une espérance ≤ 0 , et c'est une sousmartingale ssi les Y_j ont toutes une espérance ≥ 0 .

Exercice n° 4.1.1 a) Si (X_n) est une surmartingale, alors on a $\mathbb{E}(X_m | \mathcal{F}_n) \leq X_n$ p.s. pour tous $m > n$ dans \mathbb{N} , et de plus (X_n) est une surmartingale dans sa propre filtration.

b) L'ensemble des sousmartingales forme un cône convexe, stable par $(X_n, Y_n) \mapsto X_n \vee Y_n$.

c) Si (X_n) est une sousmartingale et si ϕ est une fonction convexe croissante de \mathbb{R} dans \mathbb{R} telle que chaque $\phi \circ X_n$ est intégrable, alors $(\phi \circ X_n)$ est une sousmartingale.

Si (X_n) est une martingale et si ϕ est une fonction convexe de \mathbb{R} dans \mathbb{R} telle que chaque $\phi \circ X_n$ est intégrable, alors $(\phi \circ X_n)$ est une sousmartingale.

Exercice n° 4.1.2 Notons X la chaîne de Markov sur \mathbb{Z} de transition P définie par : $P(j, j \pm 2) = 1/2$ si $j \in 3\mathbb{Z}$, et $P(j, j \pm 1) = 1/2$ sinon. Soit Y_n le résidu de X_n modulo 5. Montrer que (Y_n) est une martingale (relativement à sa propre filtration), mais n'est pas une chaîne de Markov : $\mathbb{P}(Y_{n+1} = y | Y_n = y_n, \dots, Y_1 = y_1)$ n'est pas une fonction de (y_n, y) (considérer par exemple $\mathbb{P}(Y_3 = 2 | Y_2 = 0, Y_1 = 2)$ et $\mathbb{P}(Y_3 = 2 | Y_2 = 0, Y_1 = 1)$).

Exercice n° 4.1.3 Soient N un entier ≥ 2 , a_1, \dots, a_N des réels de somme nulle, σ une v.a. uniforme sur le groupe des permutations \mathcal{S}_N , puis pour $1 \leq n < N$: $X_n := \frac{N}{N-n} \sum_{j=1}^n a_{\sigma_j}$, et \mathcal{F}_n la tribu engendrée par $\{\sigma_1, \dots, \sigma_n\}$.

a) Trouver pour $1 \leq n \leq j \leq N$ la loi conditionnelle de σ_j sachant \mathcal{F}_{n-1} .

$$\left(\text{Indication : } \mathbb{E}[\varphi(\sigma_1, \dots, \sigma_n)] = \frac{1}{C_N^n} \sum_{A \in \mathcal{C}_N^n} \frac{1}{n!} \sum_{\sigma_1, \dots, \sigma_n \in A} \varphi(\sigma_1, \dots, \sigma_n) \right)$$

b) Expliciter (pour $n \geq 2$) $X_n - X_{n-1}$, et en déduire que (X_n) est une (\mathcal{F}_n) -martingale.

Exercice n° 4.1.4 Soient $p \in]0, 1[$, G une v.a. géométrique de paramètre p : $\mathbb{P}(G = n) = (1-p)p^n$, et \mathcal{F}_n la tribu engendrée par la v.a. $n \wedge G$, pour $n \in \mathbb{N}$.

a) Vérifier que $\{\mathcal{F}_n | n \in \mathbb{N}\}$ définit une filtration.

b) Montrer que $\mathbb{E}(1_{\{G > n\}} | \mathcal{F}_n) = p 1_{\{G \geq n\}}$, pour tout $n \in \mathbb{N}$.

c) Déduire que $\mathbb{E}(G \wedge (n+1) | \mathcal{F}_n) = G \wedge n + p 1_{\{G \geq n\}}$, pour tout $n \in \mathbb{N}$.

d) Trouver $a \in \mathbb{R}$ tel que $X_n := a \times (G \wedge n) + 1_{\{G \geq n\}}$ définisse une (\mathcal{F}_n) -martingale.

e) Calculer $\mathbb{E}((X_{n+1} - X_n)^2 | \mathcal{F}_n)$, et déduire que $Y_n := X_n^2 - a \times (G \wedge (n-1))$ définit une (\mathcal{F}_n) -martingale.

Proposition 18 Si (X_n) est une surmartingale et si T est un temps d'arrêt (relativement à la filtration (\mathcal{F}_n)), alors la "surmartingale arrêtée" $(X_{n \wedge T})$ est encore une surmartingale.

Preuve $(X_{n \wedge T})$ est intégrable, puisque sa valeur absolue est majorée par $|X_0| \vee \dots \vee |X_n|$. L'adaptation découle aussitôt des exercices 2.1.2 et 2.1.1. Enfin

$$\begin{aligned}\mathbb{E}(X_{(n+1)\wedge T} - X_{n\wedge T} | \mathcal{F}_n) &= \mathbb{E}((X_{n+1} - X_n)1_{\{T \geq n+1\}} | \mathcal{F}_n) \\ &= 1_{\{T \geq n+1\}} \mathbb{E}(X_{n+1} - X_n | \mathcal{F}_n) \leq 0. \quad \diamond\end{aligned}$$

Nous avons maintenant l'important "théorème d'arrêt" de Doob :

Théorème 8 Soient S et T deux temps d'arrêt p.s. bornés, tels que p.s. $S \leq T$. Soit (X_n) une surmartingale. Alors on a p.s. $\mathbb{E}(X_T | \mathcal{F}_S) \leq X_S$.

Preuve Soit M un entier fixe majorant p.s. le temps T . X_T est intégrable, sa valeur absolue étant majorée par $|X_0| + \dots + |X_M|$. Pour tout $A \in \mathcal{F}_S$, la proposition 18 ci-dessus entraîne :

$$\mathbb{E}[(X_T - X_S)1_A] = \sum_{m=0}^M \mathbb{E}[(X_{M\wedge T} - X_{m\wedge T})1_{A \cap \{S=m\}}] \leq 0. \quad \diamond$$

En particulier, pour toute martingale (X_n) et tout temps d'arrêt borné T , on a $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

Si T n'est pas borné : le temps d'atteinte de 1 pour la marche simple sur \mathbb{Z} fournit un contreexemple.

Exercice n° 4.1.5 Soient N un entier ≥ 2 , un réel $A > 0$, une filtration $\{\mathcal{F}_n | n \in \mathbb{N}\}$, (M_n) une (\mathcal{F}_n) -martingale, (H_n) un processus (\mathcal{F}_n) -adapté,

$T := N \wedge \min\{n \in \mathbb{N}^* | |H_n| > A\}$, et $X_n := \sum_{j=1}^{T \wedge n} H_{j-1}(M_j - M_{j-1})$ pour $n \in \mathbb{N}^*$.
Montrer que T est un (\mathcal{F}_n) -temps d'arrêt, et que (X_n) est une (\mathcal{F}_n) -martingale.

4.2 Inégalités

Plusieurs inégalités expriment que pour contrôler toutes les valeurs d'une (sur)martingale, il suffit de contrôler les valeurs à ses extrémités. Commençons par une "inégalité maximale" :

Proposition 19 Soient (X_n) une surmartingale, $N \in \mathbb{N}$, et $\lambda > 0$. Nous avons :

$$\lambda \mathbb{P}\left(\max\{|X_n| | 0 \leq n \leq N\} \geq \lambda\right) \leq \mathbb{E}(X_0) + 2\mathbb{E}(X_N^-).$$

Preuve Considérons le temps d'arrêt $T := N \wedge \min\{n \in \mathbb{N} | X_n \geq \lambda\}$. Le théorème d'arrêt 8 ci-dessus nous donne (en séparant $\{\min\{n \in \mathbb{N} | X_n \geq \lambda\} \leq N\}$ de son complémentaire) :

$$\mathbb{E}(X_0) \geq \mathbb{E}(X_T) \geq \lambda \mathbb{P}\left(\max\{X_n | 0 \leq n \leq N\} \geq \lambda\right) + \mathbb{E}\left(X_N 1_{\{\max\{X_n | 0 \leq n \leq N\} < \lambda\}}\right),$$

d'où

$$\lambda \mathbb{P}\left(\max\{X_n | 0 \leq n \leq N\} \geq \lambda\right) \leq \mathbb{E}(X_0) + \mathbb{E}(X_N^-).$$

Considérons ensuite le temps d'arrêt $T' := N \wedge \min\{n \in \mathbb{N} \mid X_n \leq -\lambda\}$. Le théorème d'arrêt 8 ci-dessus nous donne :

$$\mathbb{E}(X_N) \leq \mathbb{E}(X_{T'}) \leq -\lambda \mathbb{P}\left(\min\{X_n \mid 0 \leq n \leq N\} \leq -\lambda\right) + \mathbb{E}\left(X_N 1_{\{\min\{X_n \mid 0 \leq n \leq N\} > -\lambda\}}\right),$$

d'où

$$\lambda \mathbb{P}\left(\min\{X_n \mid 0 \leq n \leq N\} \leq -\lambda\right) \leq \mathbb{E}\left(-X_N 1_{\{\min\{X_n \mid 0 \leq n \leq N\} \leq -\lambda\}}\right) \leq \mathbb{E}(X_N^-). \diamond$$

Corollaire 7 Soient (X_n) une martingale, $N \in \mathbb{N}$, $p \geq 1$, et $\lambda > 0$. Nous avons :

$$\lambda^p \mathbb{P}\left(\max\{|X_n| \mid 0 \leq n \leq N\} \geq \lambda\right) \leq \mathbb{E}(|X_N|^p).$$

Preuve L'exercice 4.1.1.c assure que $(-|X_n|^p)$ est une surmartingale, et il suffit de lui appliquer la dernière inégalité de la preuve de la proposition 19 ci-dessus, avec λ^p au lieu de λ . \diamond

Le théorème suivant énonce l'importante "inégalité de Doob".

Théorème 9 Soient (X_n) une martingale, $N \in \mathbb{N}$, et $p > 1$. Nous avons :

$$\left\| \max\{|X_n| \mid 0 \leq n \leq N\} \right\|_p \leq \frac{p}{p-1} \|X_N\|_p.$$

Preuve Posons $X_N^* := \max\{|X_n| \mid 0 \leq n \leq N\}$. La dernière ligne de la preuve de la proposition 19 ci-dessus appliquée à $(-|X_n|)$ montre que pour tout $\lambda > 0$

$$\lambda \mathbb{P}(X_N^* \geq \lambda) \leq \mathbb{E}(|X_N| 1_{\{X_N^* \geq \lambda\}}).$$

Ceci et l'inégalité de Hölder entraînent :

$$\begin{aligned} \mathbb{E}(|X_N^*|^p) &= \int_0^\infty \mathbb{P}(X_N^* \geq \lambda) d\lambda^p \leq p \int_0^\infty \mathbb{E}(|X_N| 1_{\{X_N^* \geq \lambda\}}) \lambda^{p-2} d\lambda = p \mathbb{E}\left(|X_N| \int_0^{X_N^*} \lambda^{p-2} d\lambda\right) \\ &= \frac{p}{p-1} \mathbb{E}\left(|X_N| \times |X_N^*|^{p-1}\right) \leq \frac{p}{p-1} \|X_N\|_p \times \left\| |X_N^*|^{p-1} \right\|_{\frac{p}{p-1}} = \frac{p}{p-1} \|X_N\|_p \times \|X_N^*\|_p^{p-1}, \end{aligned}$$

d'où le résultat en divisant par $\|X_N^*\|_p^{p-1}$, lorsque $X_N^* \in L^p$. Or le résultat est trivial si $X_N \notin L^p$, et $\|X_N^*\|_p \leq \|X_0\|_p + \dots + \|X_N\|_p \leq (N+1)\|X_N\|_p$ est fini si $X_N \in L^p$. \diamond

Faisant tendre N vers l'infini, on déduit aussitôt par convergence monotone :

Corollaire 8 Soient (X_n) une martingale, et $p > 1$. Nous avons :

$$\left\| \sup\{|X_n| \mid n \in \mathbb{N}\} \right\|_p \leq \frac{p}{p-1} \sup_n \|X_n\|_p.$$

Fixons deux réels a, b tels que $a < b$, et un processus (X_n) .

Soit (T_k^X) la suite croissante de temps d'arrêt définie par : $T_0^X := 0$, et pour tout $k \geq 1$:

$$T_{2k-1}^X := \min\{n \geq T_{2k-2}^X \mid X_n \leq a\}, \quad T_{2k}^X := \min\{n > T_{2k-1}^X \mid X_n \geq b\}.$$

Le "nombre de montées" de (X_n) à travers l'intervalle $[a, b]$ avant $N \in \mathbb{N}$ est :

$$M_N^{a,b}(X) := \max\{k \in \mathbb{N} \mid T_{2k}^X \leq N\}.$$

Proposition 20 Soient (X_n) une surmartingale, 2 réels $a < b$, et $N \in \mathbb{N}$. Nous avons :

$$(b - a) \mathbb{E}[M_N^{a,b}(X)] \leq \mathbb{E}[(X_N - a)^-].$$

Preuve Le théorème d'arrêt 8 nous donne pour tout $k \in \mathbb{N}^*$:

$$\begin{aligned} (b - a) \mathbb{P}(M_N^{a,b}(X) \geq k) &= (b - a) \mathbb{P}(T_{2k}^X \leq N) \leq \mathbb{E}\left[\left(X_{N \wedge T_{2k}^X} - X_{N \wedge T_{2k-1}^X}\right) 1_{\{T_{2k}^X \leq N\}}\right] \\ &= \mathbb{E}\left[\left(X_{N \wedge T_{2k}^X} - X_{N \wedge T_{2k-1}^X}\right) \left(1_{\{T_{2k-1}^X \leq N\}} - 1_{\{T_{2k-1}^X \leq N < T_{2k}^X\}}\right)\right] \\ &= \mathbb{E}\left[\left(X_{N \wedge T_{2k}^X} - X_{N \wedge T_{2k-1}^X}\right) 1_{\{T_{2k-1}^X \leq N\}}\right] + \mathbb{E}\left[\left(X_{T_{2k-1}^X} - X_N\right) 1_{\{T_{2k-1}^X \leq N < T_{2k}^X\}}\right] \\ &\leq 0 + \mathbb{E}\left[\left(a - X_N\right) 1_{\{T_{2k-1}^X \leq N < T_{2k}^X\}}\right] \leq \mathbb{E}\left[\left(a - X_N\right)^+ 1_{\{T_{2k-2}^X \leq N < T_{2k}^X\}}\right], \end{aligned}$$

d'où

$$(b - a) \mathbb{E}[M_N^{a,b}(X)] = \sum_{k \in \mathbb{N}^*} (b - a) \mathbb{P}(M_N^{a,b}(X) \geq k) \leq \mathbb{E}[(a - X_N)^+] = \mathbb{E}[(X_N - a)^-]. \quad \diamond$$

4.3 Convergence

Théorème 10 Soit (X_n) une surmartingale, telle que $\sup_n \mathbb{E}(X_n^-) < \infty$. Alors (X_n) est bornée dans L^1 , et converge presque sûrement vers une variable intégrable X_∞ .

Preuve La première assertion est très simple : on a pour tout $n \in \mathbb{N}$

$$\mathbb{E}(|X_n|) = \mathbb{E}(X_n) + 2\mathbb{E}(X_n^-) \leq \mathbb{E}(X_0) + 2 \sup_n \mathbb{E}(X_n^-) < \infty.$$

Ensuite

$$\{X. \longrightarrow X_\infty\}^c \subset \bigcup_{a,b \in \mathbb{Q}} \{\liminf_n X_n < a < b < \limsup_n X_n\} \subset \bigcup_{a,b \in \mathbb{Q}, a < b} \{M_\infty^{a,b}(X) = \infty\}.$$

Or la proposition 20 ci-dessus et le théorème de convergence monotone nous donnent :

$$\mathbb{E}[M_\infty^{a,b}(X)] \leq (b - a)^{-1} \sup_N \mathbb{E}[(X_N - a)^-] \leq (b - a)^{-1} \sup_N (\mathbb{E}(|X_N|) + |a|) < \infty$$

dès que $a < b$. Donc $\mathbb{P}(M_\infty^{a,b}(X) = \infty) = 0$, et donc $\mathbb{P}(X. \text{ n'a pas de limite dans } \overline{\mathbb{R}}) = 0$. Et le lemme de Fatou assure que $\mathbb{E}(|X_\infty|) \leq \liminf_n \mathbb{E}(|X_n|) \leq \sup_n \mathbb{E}(|X_n|) < \infty$. \diamond

Corollaire 9 Toute surmartingale positive $(X_n) \geq 0$ converge presque sûrement vers une variable intégrable X_∞ , et nous avons p.s. pour tout $n \in \mathbb{N}$: $\mathbb{E}(X_\infty | \mathcal{F}_n) \leq X_n$.

Preuve Il suffit d'appliquer le théorème 10 précédent, puis le lemme de Fatou :

$$X_n \geq \liminf_{m \rightarrow \infty} \mathbb{E}(X_{n+m} | \mathcal{F}_n) \geq \mathbb{E}(X_\infty | \mathcal{F}_n) . \diamond$$

Corollaire 10 Soient $p > 1$ et (X_n) une martingale bornée dans L^p . Alors (X_n) converge presque sûrement et dans L^p , et on a pour tout $n \in \mathbb{N}$: $\mathbb{E}(X_\infty | \mathcal{F}_n) = X_n$.

Preuve Puisque $\|X_n^-\|_1 \leq \|X_n\|_1 \leq \|X_n\|_p$, le théorème 10 ci-dessus assure la convergence presque sûre. L'inégalité de Doob (corollaire 8) montre que $\sup_n |X_n| \in L^p$, et donc que $X_\infty \in L^p$. Puisque $|X_n - X_\infty|^p \leq (2 \sup_n |X_n|)^p$, on a bien $\|X_n - X_\infty\|_p \rightarrow 0$ par convergence dominée. Enfin l'inégalité de Jensen fournit :

$$\|\mathbb{E}(X_\infty | \mathcal{F}_n) - X_n\|_p^p \leq \limsup_m \mathbb{E}[|\mathbb{E}(X_\infty - X_{n+m} | \mathcal{F}_n)|^p] \leq \limsup_m \|X_\infty - X_{n+m}\|_p^p = 0 . \diamond$$

Exercice n° 4.3.0 Soient $(S_n = X_1 + \dots + X_n)$ une marche aléatoire, où X_1 est une variable aléatoire de Bernoulli, et soit $M_n := e^{S_n} / \mathbb{E}(e^{S_n})$. Vérifier que M est une martingale bornée dans L^1 , qui converge presque sûrement vers 0. Comparer avec le Corollaire 10.

Exercice n° 4.3.1 Soient $p > 1$, et Y une v.a.r. appartenant à $L^p(\mathcal{F}_\infty, \mathbb{P})$.

a) Montrer que $X_n := \mathbb{E}(Y | \mathcal{F}_n)$ définit une martingale, convergeant presque sûrement et dans L^p vers Y . (Vérifier que si $Z \in L^1(\mathcal{F}_\infty, \mathbb{P})$ et $\mathbb{E}(Z | \mathcal{F}_n) = 0 \ \forall n$, alors $Z = 0$ p.s., en considérant successivement : $\mathcal{P} := \{A \in \mathcal{F}_\infty | \mathbb{E}(Z 1_A) = 0\}$; une partie maximale \mathcal{E} stable par réunions et intersections finies, contenue dans \mathcal{P} et contenant $\bigcup_n \mathcal{F}_n$; puis, pour toute $A := \bigcup_n A_n$ union dénombrable croissante de $(A_n) \subset \mathcal{E}$: $\mathcal{A} := \{(B \cap A) \cup B' | B, B' \in \mathcal{E}\}$; idem pour toute intersection dénombrable décroissante; et enfin, $\mathcal{E}' := \{A \in \mathcal{E} | A^c \in \mathcal{E}\}$.)

b) Qu'obtient-on si on suppose $Y \in L^p(\mathcal{F}, \mathbb{P})$ (i.e. Y peut n'être pas \mathcal{F}_∞ -mesurable) ?

Soient $\{Z_n | n \in \mathbb{N}^*\}$ une suite de v.a.r. indépendantes, \mathcal{F}_n la tribu engendrée par $\{Z_k | k \leq n\}$, \mathcal{G}_n la tribu engendrée par $\{Z_k | k > n\}$, et $\mathcal{G}_\infty := \bigcap_n \mathcal{G}_n$.

c) Dédurre de (a) (avec $Y = 1_A$) la loi du 0-1 : tout $A \in \mathcal{G}_\infty$ est de probabilité 0 ou 1.

Exercice n° 4.3.2 Une urne contient initialement n boules noires et b boules blanches. On tire une boule dans l'urne, selon la loi uniforme, puis on remet la boule tirée avec a autres boules de la même couleur. On itère indéfiniment cette procédure. Notons X_ℓ la proportion de boules noires dans l'urne après le ℓ -ième tirage (de sorte que $X_0 = n/(n+b)$).

a) Montrer que (X_ℓ) est une chaîne de Markov inhomogène sur $\mathbb{Q} \cap [0, 1]$.

b) Montrer que (X_ℓ) est l'image par une application à expliciter d'une chaîne de Markov homogène sur \mathbb{N}^2 , dont on précisera le noyau de transition.

c) Montrer que (X_ℓ) est une martingale, qui converge p.s. et dans tous les L^p .

d) Dans le cas $n = b = a = 1$, calculer la loi de (X_ℓ) , et en déduire la loi de X_∞ .

Exercice n° 4.3.3 a) Soit $\{B_n \mid n \in \mathbb{N}^*\}$ une suite i.i.d. de v.a. de Bernoulli prenant les valeurs 0, 2 avec probabilité 1/2. Posons $X_0 = 1$, puis $X_n := \prod_{k=1}^n B_k$, pour tout $n \in \mathbb{N}^*$.

Montrer que (X_n) est une chaîne de Markov (préciser sa transition) et une martingale p.s. convergente, mais non convergente dans les L^p , et telle que $\mathbb{E}(X_\infty \mid \mathcal{F}_n) \neq X_n$.

b) Reprenons les mêmes B, X , sauf que cette fois les 2 valeurs prises par les B_n sont ± 1 . Montrer que (X_n) est une chaîne de Markov centrée et n'est pas une martingale p.s. convergente, bien que $|X_n| = 1$ pour tout n .

Exercice n° 4.3.4 Soit $\{X_n \mid n \in \mathbb{N}^*\}$ une suite de v.a.r. indépendantes, centrées et de carré intégrable. Pour tout $n \in \mathbb{N}^*$, notons : $M_n := \sum_{j=1}^n X_j$, $A_n := \sum_{j=1}^n \text{Var}(X_j)$, $N_n := M_n^2 - A_n$, et \mathcal{F}_n la tribu engendrée par $\{X_1, \dots, X_n\}$. ($M_0 = A_0 = 0$, et $\mathcal{F}_0 = \{\emptyset, \Omega\}$).

a) Que peut-on dire du processus (M_n) ?

b) Calculer $\mathbb{E}((M_{n+1} - M_n)^2 \mid \mathcal{F}_n)$ de deux façons différentes, et déduire que (N_n) est une (\mathcal{F}_n) -martingale.

c) Déduire que si $A_\infty < \infty$, alors (N_n) et (M_n) convergent presque sûrement.

Réciproquement, supposons dorénavant que $\mathbb{P}[(M_n) \text{ converge}] > 0$, et pour $b \in \mathbb{R}_+^*$ posons $T_b := \min\{n \in \mathbb{N} \mid |M_n| > b\}$.

d) Montrer que T_b est un temps d'arrêt, et qu'on peut prendre b tel que $\mathbb{P}(T_b = \infty) > 0$.

e) Déduire de (b) que $\mathbb{E}(A_{T_b \wedge n}) \leq \mathbb{E}[(|X_{T_b}|1_{\{T_b < \infty\}} + b)^2]$ pour tout $n \in \mathbb{N}$.

f) Déduire que si (X_n) est uniformément bornée, alors $A_\infty < \infty$.

Soient (a_n) une suite réelle, et (B_n) une suite i.i.d. de v.a. de Bernoulli prenant les valeurs ± 1 avec probabilité 1/2, 1/2.

g) Montrer que $\sum_n B_n a_n$ converge p.s. ssi $\sum_n a_n^2$ converge.

h) Montrer que si $\sum_n a_n^2$ diverge, alors p.s. $\liminf_k \sum_{n \leq k} B_n a_n < \limsup_k \sum_{n \leq k} B_n a_n$.

Exercice n° 4.3.5 Soient $N \in \mathbb{N}^*$, et $\{B_n \mid n \in \mathbb{N}^*\}$ une suite i.i.d. de v.a. de Bernoulli prenant les valeurs 1, -1 avec probabilité $p, 1-p$ respectivement, p étant fixé dans $] \frac{1}{2}, 1[$.

Posons $\alpha := p \log p + (1-p) \log(1-p) + \log 2$.

Considérons un jeu se déroulant en N parties, et tel qu'à la n -ième partie on gagne $B_n Y_n$ si on a misé Y_n . Notons \mathcal{F}_n la tribu engendrée par $\{B_1, \dots, B_n\}$, et G_n la fortune du joueur après la n -ième partie, pour tout $n \in \mathbb{N}$. G_0 est une constante positive donnée.

Le processus (Y_n) représente la stratégie du joueur, et doit donc être "prévisible" : Y_n doit être \mathcal{F}_{n-1} -mesurable, pour tout $n \in \mathbb{N}$. De plus, Y_n doit appartenir à $[0, G_{n-1}[$.

a) Montrer que, pour toute stratégie, $(\log G_n - \alpha n)$ est une surmartingale.

b) Montrer que, pour une certaine stratégie, $(\log G_n - \alpha n)$ est une martingale.

c) Quelle est la stratégie maximisant $\mathbb{E}[\log(G_N/G_0)]$?

Exercice n° 4.3.6 (Décomposition de Doob) : Tout processus adapté $X = \{X_n \mid n \in \mathbb{N}\}$ de v.a.r. intégrables se décompose de façon unique en $X = X_0 + M + A$, où M est une martingale nulle en 0 et A est un processus nul en 0, intégrable et “prévisible” (A_n est \mathcal{F}_{n-1} -mesurable pour tout $n \in \mathbb{N}^*$).

a) Montrer que nécessairement $A_n = \sum_{k=1}^n \mathbb{E}(X_k - X_{k-1} \mid \mathcal{F}_{k-1})$ p.s..

b) Prouver l'énoncé ci-dessus.

c) Montrer que X est une surmartingale ssi A est p.s. décroissant.

Exercice n° 4.3.7 Fixons un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et une filtration “renversée” $\{\mathcal{F}_n \mid n \in -\mathbb{N}\} : \mathcal{F}_{-k-1} \subset \mathcal{F}_{-k} \subset \mathcal{F}$ pour tout $k \in \mathbb{N}$. Soient $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ et pour tout $k \in \mathbb{N} : X_{-k} := \mathbb{E}(Y \mid \mathcal{F}_{-k})$.

a) Vérifier que $X = \{X_n \mid n \in -\mathbb{N}\}$ est une (\mathcal{F}_n) -martingale bornée dans L^1 .

b) Vérifier que la preuve du théorème 11 de convergence (de certaines surmartingales) s'adapte (sans difficulté) pour montrer que X converge p.s., vers une certaine $X_{-\infty} \in L^1$.

c) Montrer : $\mathbb{E}(|X_n| 1_{\{|X_n| \geq c\}}) \leq \mathbb{E}(|Y| 1_{\{|X_n| \geq c\}})$, et $\lim_{c \rightarrow \infty} \sup_{n \in -\mathbb{N}} \mathbb{E}(|X_n| 1_{\{|X_n| \geq c\}}) = 0$.

d) Montrer que X converge dans L^1 et que $X_{-\infty} = \mathbb{E}(Y \mid \bigcap_n \mathcal{F}_n)$ presque sûrement.

Exercice n° 4.3.8 *Preuve martingale de la loi forte des grands nombres.*

Soit $\{X_n \mid n \in \mathbb{N}^*\}$ une suite de v.a.r.i.i.d. intégrables, d'espérance m . Pour tout $n \in \mathbb{N}^*$, notons : $S_n := \sum_{j=1}^n X_j$, et \mathcal{G}_n la tribu engendrée par $\{S_k \mid k \geq n\}$.

a) Montrer que $\mathbb{E}(X_1 \mid \mathcal{G}_n) = \mathbb{E}(X_1 \mid S_n)$, pour tout $n \in \mathbb{N}^*$. par b) Montrer que $\mathbb{E}(X_1 \mid S_n) = S_n/n$, pour tout $n \in \mathbb{N}^*$.

c) Dédurre de l'exercice précédent que S_n/n converge p.s. et dans L^1 vers une v.a. Λ .

d) Dédurre de l'exercice 4.3.1 (loi du 0-1) que Λ est p.s. constante. Conclure.

Exercice n° 4.3.9 Soit (X, P) une chaîne de Markov irréductible récurrente, ν une mesure invariante non nulle pour P , et μ une mesure P -excessive : $\mu P \leq \mu$. Soient f définie sur E par $f(e) := \mu_e / \nu_e$, et $Q(Q(x, y) := P(y, x) \times \nu_y / \nu_x)$ le noyau adjoint de P (correspondant à la chaîne retournée, voir l'exercice 3.4.1). Montrer successivement que : $Qf \leq f$, $f \circ X$ est une surmartingale (pour quelle(s) probabilité(s) ?), f est constante, ν est unique à une constante multiplicative près.

5 Processus de Poisson

Définition 17 On appelle *processus de comptage* tout processus $(N_t \mid t \in \mathbb{R}_+)$ donné par $N_t := \text{Card}(\mathcal{E} \cap [0, t])$, où la donnée \mathcal{E} est une partie discrète aléatoire de \mathbb{R}_+ .

Le processus de comptage N est dit à *accroissements indépendants* lorsque pour tout n dans

\mathbb{N}^* et tous $s_1 < t_1 \leq \dots \leq s_n < t_n$, les variables $(N_{t_1} - N_{s_1}), \dots, (N_{t_n} - N_{s_n})$ sont indépendantes. Cela signifie l'indépendance des nombres de points de \mathcal{E} situés dans des intervalles disjoints deux à deux.

Remarques 1) Un processus de comptage est nécessairement croissant, à valeurs dans $\overline{\mathbb{N}}$, et presque sûrement continu à droite en tout temps (y compris aléatoire) :

$$\mathbb{P}(\inf N_{S_n} \neq N_S) \leq \mathbb{P}(N_{S_n} - N_S \geq 1) = \mathbb{P}(\mathcal{E} \cap]S, S_n] \neq \emptyset) \searrow 0 \text{ si } S_n \searrow S \text{ p.s.}$$

2) Les points de \mathcal{E} sont généralement des instants successifs de survenue d'événements (d'un type donné a priori : arrivées dans une file, passages de bus, désintégrations d'atomes, etc.), qu'on observe au cours du temps t . Pour tous $0 \leq s < t$, la variable aléatoire $N_t - N_s$ décompte le nombre de ces événements survenus dans l'intervalle de temps $]s, t]$.

5.1 Processus de Poisson homogènes en dimension 1

Définition 18 On appelle *processus de Poisson (homogène) d'intensité λ* (quelconque dans \mathbb{R}_+^*) tout processus de comptage N à accroissements indépendants nul en 0 et tel que pour tous $s, t \in \mathbb{R}_+$ la loi de $(N_{s+t} - N_s)$ est $\mathcal{P}(\lambda t)$, la loi de Poisson de paramètre λt .

Fixons dorénavant un processus de Poisson N , de filtration (\mathcal{F}_t) et d'intensité λ .

Exercice n° 5.1.1 a) Un processus de Poisson est à accroissements stationnaires : la loi de tout vecteur $(N_{a+t_1} - N_{a+s_1}, \dots, N_{a+t_n} - N_{a+s_n})$ est indépendante de $a \in \mathbb{R}_+$.

b) Un processus de Poisson est p.s. fini à tout instant : $\mathbb{P}\left(\bigcap_t \{N_t < \infty\}\right) = 1$: les événements qu'il décompte ne peuvent pas s'accumuler en un instant fini.

c) En temps t petit, on a : $\mathbb{P}(N_t = 1) = \lambda t + o(t)$ et $\mathbb{P}(N_t \geq 1) = \lambda t + o(t)$.

d) On a presque sûrement : $\lim_{t \rightarrow \infty} N_t/t = \lambda = \mathbb{E}(N_t/t)$.

Exercice n° 5.1.2 a) Pour tout processus de Poisson N d'intensité λ , tout temps d'arrêt p.s. fini S de la filtration (\mathcal{F}_t) engendrée par (N_t) , et tout $t > 0$, la variable aléatoire $(N_{S+t} - N_S)$ est indépendante de \mathcal{F}_S et a pour loi $\mathcal{P}(\lambda t)$.

(commencer par le cas de S à valeurs discrètes, puis l'approcher par une suite décroissante.)

b) $(N_{S+t} - N_S \mid t \in \mathbb{R}_+)$ est un processus de Poisson d'intensité λ , indépendant de \mathcal{F}_S .

Exercice n° 5.1.3 Montrer que $S_1 := \inf\{t > 0 \mid N_t > 0\}$ est un temps d'arrêt presque sûrement fini et > 0 .

Lemme 7 On a presque sûrement $N_{S_1} = 1$ (où $S_1 := \inf\{t > 0 \mid N_t > 0\}$).

Preuve On a $N_{S_1} \geq 1$ par continuité à droite, et pour tout $\varepsilon > 0$:

$$\mathbb{P}(N_{S_1} \geq 2) = \sum_{k \in \mathbb{N}} \mathbb{P}(N_{S_1} \geq 2, k\varepsilon < S_1 \leq (k+1)\varepsilon) \leq \sum_{k \in \mathbb{N}} \mathbb{P}(N_{(k+1)\varepsilon} - N_{k\varepsilon} \geq 2, N_{k\varepsilon} = 0)$$

$$= \sum_{k \in \mathbb{N}} \mathbb{P}(N_\varepsilon \geq 2) \times \mathbb{P}(N_{k\varepsilon} = 0) \leq (\lambda\varepsilon)^2 \times (1 - e^{-\lambda\varepsilon})^{-1} \sim \lambda\varepsilon . \diamond$$

Corollaire 11 *Les points de \mathcal{E} s'ordonnent en une suite strictement croissante de variables aléatoires tendant vers $+\infty$: $S_0 := 0 < S_1 < \dots < S_n < \dots$. Pour tout $n \in \mathbb{N}$, S_n est le n -ième point de \mathcal{E} , et posant $(N \circ \theta^S)_t := N_{S+t} - N_S$, nous avons :*

$$S_{n+1} = S_n + S_1 \circ \theta^{S_n} = S_1 + S_n \circ \theta^{S_1}.$$

Preuve Définissons par récurrence $S_0 := 0$ et $S_{n+1} := \inf\{t > 0 \mid N_t > N_{S_n}\}$. Les exercices (5.1.3), (5.1.2.b) et le lemme 7 assurent par récurrence que ce sont des temps d'arrêt p.s. finis de la filtration engendrée par (N_t) , et que $\mathbb{P}(N_{S_{n+1}} - N_{S_n} = 1) = 1$. Ceci montre que les points de \mathcal{E} surviennent bien lors d'une suite strictement croissante, et que S_n est le n -ième point de \mathcal{E} . L'exercice (5.1.1.b) montre que cette suite tend vers l'infini presque sûrement. Le reste est comme pour les marches aléatoires. \diamond

Théorème 11 *Pour tout $n \in \mathbb{N}^*$, notons $T_n := S_n - S_{n-1}$ l'intervalle de temps entre les $(n-1)$ -ième et n -ième points de \mathcal{E} .*

La suite $\{T_n, n \in \mathbb{N}^\}$ est i.i.d., de loi commune la loi $\mathcal{E}(\lambda)$ exponentielle de paramètre λ .*

Preuve Fixons $n \in \mathbb{N}^*$, et $0 \leq s_1 < t_1 \leq \dots \leq s_n < t_n$. Notons \mathcal{L}_n la loi de (S_1, \dots, S_n) , et μ_n la loi de densité $\lambda^n 1_{\{\sigma_1 < \dots < \sigma_n\}} e^{-\lambda\sigma_n}$ par rapport à la mesure de Lebesgue $d\sigma = d\sigma_1 \dots d\sigma_n$ sur \mathbb{R}_+^n . Nous avons

$$\begin{aligned} \mathcal{L}_n([s_1, t_1] \times \dots \times]s_n, t_n]) &= \mathbb{P}(N_{s_1} = 0, N_{t_1} - N_{s_1} = 1, \dots, N_{s_n} - N_{t_{n-1}} = 0, N_{t_n} - N_{s_n} \geq 1) \\ &= \mathbb{P}(N_{s_1} = 0) \times \mathbb{P}(N_{t_1 - s_1} = 1) \times \dots \times \mathbb{P}(N_{s_n - t_{n-1}} = 0) \times \mathbb{P}(N_{t_n - s_n} \geq 1) \\ &= \lambda^{n-1} e^{-\lambda s_n} (t_1 - s_1) \dots (t_{n-1} - s_{n-1}) (1 - e^{-\lambda(t_n - s_n)}) = \mu_n([s_1, t_1] \times \dots \times]s_n, t_n]), \end{aligned}$$

montrant que les deux lois \mathcal{L}_n et μ_n sont égales sur les produits d'intervalles disjoints deux à deux. Ceci entraîne que les lois marginales de \mathcal{L}_n ne possèdent pas d'atome, comme celles de μ_n . Comme les mesures boréliennes \mathcal{L}_n et μ_n sont supportées par

$\{0 < \sigma_1 < \dots < \sigma_n\}$, nous avons donc (par convergence dominée) pour un pavé quelconque

$$A := \prod_{k=1}^n]u_k, v_k] \subset \mathbb{R}_+^n :$$

$$\begin{aligned} \mathcal{L}_n(A) &= \lim_{\varepsilon \searrow 0} \sum_{0 \leq j_1 < \dots < j_n} 1_{\bigcap_{k=1}^n \{u_k < j_k \varepsilon < v_k\}} \mathcal{L}_n\left(\prod_{k=1}^n]j_k \varepsilon, (j_k + 1)\varepsilon\right) \\ &= \lim_{\varepsilon \searrow 0} \sum_{0 \leq j_1 < \dots < j_n} 1_{\bigcap_{k=1}^n \{u_k < j_k \varepsilon < v_k\}} \mu_n\left(\prod_{k=1}^n]j_k \varepsilon, (j_k + 1)\varepsilon\right) = \mu_n(A). \end{aligned}$$

Donc les mesures boréliennes \mathcal{L}_n et μ_n sont égales. Nous en déduisons par le changement de variables $\sigma_k = t_1 + \dots + t_k$, $1 \leq k \leq n$, de jacobien égal à 1, pour toute fonction borélienne positive F sur \mathbb{R}_+^n :

$$\begin{aligned} \mathbb{E}[F(T_1, \dots, T_n)] &= \mathbb{E}[F(S_1, \dots, S_n - S_{n-1})] = \int F(\sigma_1, \dots, \sigma_n - \sigma_{n-1}) d\mathcal{L}_n(\sigma) \\ &= \int_{0 < \sigma_1 < \dots < \sigma_n} F(\sigma_1, \dots, \sigma_n - \sigma_{n-1}) \lambda^n e^{-\lambda \sigma_n} d\sigma_1 \dots d\sigma_n \\ &= \lambda^n \int_{\mathbb{R}_+^n} F(t_1, \dots, t_n) e^{-\lambda(t_1 + \dots + t_n)} dt_1 \dots dt_n = \int F(t_1, \dots, t_n) \prod_{k=1}^n 1_{\{t_k > 0\}} \lambda e^{-\lambda t_k} dt_k. \quad \diamond \end{aligned}$$

Exercice 5.1.4 a) Montrer que pour tout $n \in \mathbb{N}^*$ la loi de S_n est la loi Gamma de paramètres n et λ , i.e. la loi de densité sur \mathbb{R}_+ : $s \rightarrow \frac{(\lambda s)^{n-1}}{(n-1)!} \times \lambda e^{-\lambda s}$.

b) Pour $t > 0$ et $f \in L^1([0, t])$, calculer $\mathbb{E}\left(\sum_{n \in \mathbb{N}^*} f(S_n) 1_{\{S_n \leq t\}}\right)$.

c) Fixons $t > 0$ et $n \in \mathbb{N}^*$. Montrer que conditionnellement par rapport à $\{N_t = n\}$, le vecteur (S_1, \dots, S_n) admet la densité de Dirichlet : $(s_1, \dots, s_n) \rightarrow \frac{n!}{t^n} 1_{\{0 < s_1 < \dots < s_n < t\}}$.

d) Montrer que (pour $t > 0$ et $n \in \mathbb{N}^*$) si U_1, \dots, U_n sont n variables i.i.d. uniformes sur $[0, t]$, et si (V_1, \dots, V_n) est le vecteur obtenu en ordonnant U_1, \dots, U_n dans le sens croissant, alors la "statistique d'ordre" (V_1, \dots, V_n) admet aussi la densité de Dirichlet ci-dessus.

e) Dédire que pour $0 < s < t$, conditionnellement par rapport à $\{N_t = n\}$, la loi de N_s est $\mathcal{B}(n, s/t)$.

f) Dédire de c), d) que pour $0 \leq s < t$, conditionnellement par rapport à $\{N_t - N_s = n\}$, les n points de $\mathcal{E} \cap]s, t]$ sont des v.a.i.i.d. de loi uniforme $\mathcal{U}[s, t]$ sur $[s, t]$.

Il est frappant que la réciproque du théorème 11 soit exacte, permettant par exemple de simuler un processus de Poisson.

Théorème 12 Soit $\{T_n; n \in \mathbb{N}^*\}$ une suite i.i.d. de variables exponentielles de paramètre $\lambda > 0$. Soit (pour tout $t \in \mathbb{R}_+$) $N_t := \sum_{n \in \mathbb{N}^*} 1_{\{T_1 + \dots + T_n \leq t\}}$. Alors N est un processus de Poisson d'intensité λ .

Preuve Fixons $n \in \mathbb{N}^*$, $k_1, \dots, k_n \in \mathbb{N}$, et $0 = t_0 < t_1 < \dots < t_n$. Nous devons établir que l'événement $\bigcap_{j=1}^n \{N_{t_j} - N_{t_{j-1}} = k_j\}$ a la probabilité voulue par la définition 18. Or

$$\alpha_n := \mathbb{P}\left(\bigcap_{j=1}^n \{N_{t_j} - N_{t_{j-1}} = k_j\}\right) = \mathbb{P}\left(\bigcap_{j=1}^n \{N_{t_j} = \sigma_j\}\right),$$

où on a posé $\sigma_j := k_1 + \dots + k_j$ pour tout $j \in \{1, \dots, n\}$.

Posons aussi $\sigma_0 := S_0 := 0$ et $S_m := T_1 + \dots + T_m$ pour tout $m \in \mathbb{N}^*$.

Le changement de variables de la preuve du théorème 11 montre aussitôt que le vecteur (S_1, \dots, S_m) admet la densité $\lambda^m 1_{\{s_1 < \dots < s_m\}} e^{-\lambda s_m}$ par rapport à la mesure de Lebesgue $ds = ds_1 \dots ds_m$ sur \mathbb{R}_+^m . De plus on a par définition $N_{t_j} = \sigma_j \iff S_{\sigma_j} \leq t_j < S_{1+\sigma_j}$.

Nous avons donc

$$\begin{aligned}
\alpha_n &= \mathbb{P}\left(\bigcap_{j=1}^n \{S_{\sigma_j} \leq t_j < S_{1+\sigma_j}\}\right) = \lambda^{1+\sigma_n} \int_{\{0 < s_1 < \dots < s_{1+\sigma_n}\}} e^{-\lambda s_{1+\sigma_n}} \prod_{j=1}^n 1_{\{s_{\sigma_j} \leq t_j < s_{1+\sigma_j}\}} ds_1 \dots ds_{1+\sigma_n} \\
&= \lambda^{1+\sigma_n} \int_{\{0 < s_1 < \dots < s_{\sigma_1} < t_1 < s_{1+\sigma_1} < \dots < s_{\sigma_2} < t_2 < \dots < s_{\sigma_n} < t_n < s_{1+\sigma_n}\}} e^{-\lambda s_{1+\sigma_n}} ds_1 \dots ds_{1+\sigma_n} \\
&= \lambda^{\sigma_n} e^{-\lambda t_n} \int_{\{0 < s_1 < \dots < s_{\sigma_1} < t_1 < s_{1+\sigma_1} < \dots < s_{\sigma_2} < t_2 < \dots < s_{\sigma_n} < t_n\}} ds_1 \dots ds_{\sigma_n} \\
&= \lambda^{\sigma_n} e^{-\lambda t_n} \times \prod_{j=1}^n \int_{\{t_{j-1} < s_{1+\sigma_{j-1}} < \dots < s_{\sigma_j} < t_j\}} ds_{1+\sigma_{j-1}} \dots ds_{\sigma_j} \\
&= \lambda^{\sigma_n} e^{-\lambda t_n} \prod_{j=1}^n \int_{\{t_{j-1} < s_1 < \dots < s_{k_j} < t_j\}} ds_1 \dots ds_{k_j} = \lambda^{\sigma_n} e^{-\lambda t_n} \prod_{j=1}^n (t_j - t_{j-1})^{k_j} \int_{\{0 < s_1 < \dots < s_{k_j} < 1\}} ds_1 \dots ds_{k_j} \\
&= \lambda^{\sigma_n} e^{-\lambda t_n} \times \prod_{j=1}^n \frac{(t_j - t_{j-1})^{k_j}}{(k_j)!} = \prod_{j=1}^n \left(e^{-\lambda(t_j - t_{j-1})} \times \frac{(\lambda(t_j - t_{j-1}))^{k_j}}{(k_j)!} \right). \quad \diamond
\end{aligned}$$

Exercice 5.1.5 a) Soient N et N' deux processus de Poisson indépendants, de paramètres λ et λ' . Montrer que $N + N'$ définit un processus de Poisson de paramètre $\lambda + \lambda'$.

b) Soit $\{X_n \mid n \in \mathbb{N}\}$ une suite i.i.d. indépendante de N , à valeurs dans \mathbb{N} . Le processus défini par $\Sigma_t := \sum_{n=1}^{N_t} X_n$ est appelé “processus de Poisson composé”. (Exemple : N peut décrire le nombre de clients d’un magasin et X_n la dépense du n -ième client.) Calculer la transformée de Laplace de Σ_t en fonction de celle des X_n et de la fonction génératrice de N_t , puis l’espérance et la variance de Σ_t .

Exercice 5.1.6 a) Soit N un processus de Poisson d’intensité λ comptant les passages successifs de voitures sur une route. Un piéton a besoin d’une durée $d > 0$ entre deux voitures pour traverser la route. Calculer la transformée de Laplace du temps Y qu’il doit attendre avant de pouvoir traverser. (On trouve $a \mapsto \frac{a+\lambda}{ae^{\lambda d} + \lambda e^{-ad}}$.) Calculer l’espérance de Y . (On trouve $\frac{e^{\lambda d} - 1 - \lambda d}{\lambda}$.)

b) Considérons une route à deux voies, les voitures étant comptées par deux processus de Poisson indépendants, d’intensité λ dans un sens et λ' dans l’autre, et nécessitant des durées respectives d et d' pour être traversées. Soient Z le temps que le piéton doit attendre pour traverser les deux voies d’un coup, et Z' le temps qu’il doit attendre pour traverser les deux voies sachant qu’il peut les traverser séparément du fait de la présence d’un refuge au milieu de la route. Donner les transformées de Laplace de Z et de Z' , leur espérances, et discuter l’intérêt du refuge.

Proposition 21 (*Processus de Poisson marqué*) Fixons un entier $r \geq 2$ et une probabilité $p = (p_1, \dots, p_r)$ sur l'ensemble $\{1, \dots, r\}$. Soit $\{Y_n | n \in \mathbb{N}^*\}$ une suite i.i.d. indépendante de N , de loi commune p . Accolons à chaque instant S_n décompté par N le type Y_n , et notons N^j le processus de comptage des instants de type j . Alors chaque N^j est un processus de Poisson d'intensité λp_j , et les processus N^j sont indépendants.

Preuve Conditionnellement à $\{N_t = k\}$, les instants S_1, \dots, S_k se répartissent dans les différents types suivant la loi multinômiale de paramètres k et p_1, \dots, p_r , par définition de la procédure. Nous avons donc pour tous $k_1, \dots, k_r \in \mathbb{N}$, notant $k := k_1 + \dots + k_r$:

$$\mathbb{P}(N_t^1 = k_1, \dots, N_t^r = k_r) = \frac{k!}{k_1! \dots k_r!} (p_1)^{k_1} \dots (p_r)^{k_r} \times e^{-\lambda t} \frac{(\lambda t)^k}{k!} = \prod_{j=1}^r e^{-\lambda p_j t} \frac{(\lambda p_j t)^{k_j}}{k_j!}.$$

Ceci prouve que pour chaque $t \geq 0$ les variables N_t^1, \dots, N_t^r sont indépendantes et ont chacune la loi de Poisson annoncée. La même observation vaut aussi bien pour $0 \leq t \leq t'$ et $N_{t'}^1 - N_t^1, \dots, N_{t'}^r - N_t^r$. La nullité en 0 est claire. Vérifions enfin la stationnarité, et l'indépendance des accroissements, en considérant des instants $0 < t_1 < \dots < t_q$, et des entiers naturels k_j^i , et en posant $m_j := k_j^1 + \dots + k_j^r$:

$$\begin{aligned} \mathbb{P} \left[\bigcap_{j=1}^q \bigcap_{i=1}^r \{N_{t_j}^i - N_{t_{j-1}}^i = k_j^i\} \right] &= \mathbb{P} \left[\bigcap_{j=1}^q \left(\bigcap_{i=1}^r \{N_{t_j}^i - N_{t_{j-1}}^i = k_j^i\} \cap \{N_{t_j} - N_{t_{j-1}} = m_j\} \right) \right] \\ &= \prod_{j=1}^q \mathbb{P} \left(\bigcap_{i=1}^r \{N_{t_j}^i - N_{t_{j-1}}^i = k_j^i\} / N_{t_j} - N_{t_{j-1}} = m_j \right) \times \mathbb{P}(N_{t_j} - N_{t_{j-1}} = m_j) \end{aligned}$$

(puisque les événements $\bigcap_{i=1}^r \{N_{t_j}^i - N_{t_{j-1}}^i = k_j^i\} \cap \{N_{t_j} - N_{t_{j-1}} = m_j\}$ appartiennent aux q tribus indépendantes $\sigma\{N_{t_j}^i - N_{t_{j-1}}^i, Y_{m_1+\dots+m_{j-1}+1}, \dots, Y_{m_1+\dots+m_j}\}$)

$$= \prod_{j=1}^q \prod_{i=1}^r e^{-\lambda p_i (t_j - t_{j-1})} \frac{[\lambda p_i (t_j - t_{j-1})]^{k_j^i}}{(k_j^i)!} \quad (\text{comme lorsqu'il n'y a qu'un seul temps } t). \quad \diamond$$

Exercice 5.1.7 Le flux (poissonnien) des trams passant aux halles est en moyenne de 12 trams par heure et comporte une proportion $p = 40\%$ de trams de la ligne D (et de $1 - p$ pour la ligne A).

- Quelle est la probabilité qu'au moins 2 trams D passent en un quart d'heure donné ?
- Sachant que 3 trams D sont passés en 20 minutes, quel est le nombre moyen de trams passés dans le même temps ?
- Sachant que 10 trams sont passés en 45 minutes, quelle est la probabilité que la moitié soient des trams D ?

Exercice 5.1.8 Pour $t > 0$ fixé, vérifier que $S_{N_t} \leq t < S_{1+N_t}$. S_{N_t} et S_{1+N_t} sont les points de \mathcal{E} précédant et suivant directement l'instant t . Autrement dit, on peut les voir

comme les instants de naissance et de décès de l'intervalle "vivant" à l'instant t . Notons $T'_t := t - S_{N_t}$ et $T''_t := S_{1+N_t} - t$ respectivement l'âge de cet intervalle vivant et le temps qu'il lui reste à vivre.

- Montrer que pour $a, b > 0$ on a $\mathbb{P}(T'_t > a, T''_t > b) = e^{-\lambda(a+b)} 1_{\{t > a\}}$.
- Déduire que T'_t et T''_t sont indépendants et donner leur lois.
- Comparer la durée de vie moyenne de l'intervalle vivant à l'instant t avec celle du n -ième intervalle de \mathcal{E} ; explication ?

Exercice 5.1.9 On veut estimer λ par le maximum de vraisemblance $\hat{\lambda}$ dans les deux cas suivant. a) On observe jusqu'à l'instant $t > 0$ fixé, et on connaît donc N_t, S_1, \dots, S_{N_t} . Calculer $\hat{\lambda}$, et vérifier qu'il est sans biais. (Utiliser la densité sachant N_t .)

b) On observe jusqu'à l'instant S_n , pour n fixé, et on connaît donc S_1, \dots, S_n . Calculer $\hat{\lambda}$, et vérifier qu'il est biaisé.

5.2 Autres processus de Poisson

Voici la généralisation de la définition 18.

Définition 19 Soient \mathcal{E} une partie aléatoire de \mathbb{R}^d et μ une mesure positive sur \mathbb{R}^d . Pour tout borélien A de \mathbb{R}^d , notons $N(A) \in \overline{\mathbb{N}}$ le nombre des points de $\mathcal{E} \cap A$. La fonction aléatoire N est appelée *processus de Poisson ponctuel d'intensité μ sur \mathbb{R}^d* si pour tout borélien μ -intégrable A la loi de $N(A)$ est $\mathcal{P}(\mu(A))$ et si les variables $N(A)$ correspondant à des A disjoints deux à deux sont indépendantes.

Ce processus de Poisson ponctuel est dit *homogène d'intensité $\lambda > 0$* lorsque la mesure μ est égale à λ fois la mesure de Lebesgue de \mathbb{R}^d .

Exercice 5.2.1 Fixons $d \in \mathbb{N}^*$, $\lambda > 0$, et pour tout $R > 0$ notons B_R la boule $B(0; R) \subset \mathbb{R}^d$ et $V_R = \pi^{\frac{d}{2}} R^d / \Gamma(1 + \frac{d}{2})$ son volume. Considérons $r := [\lambda V_R]$ points aléatoires indépendants répartis uniformément dans B_R , formant une partie \mathcal{E}_R de B_R . Fixons $m \in \mathbb{N}^*$, m boréliens disjoints 2 à 2 : A_1, \dots, A_m inclus dans B_R , $A_0 := B_R \setminus \cup_{j=1}^m A_j$, et pour $0 \leq j \leq m$ notons $N_R(A_j)$ le nombre des points de \mathcal{E}_R contenus dans A_j .

- Donner la loi du vecteur $(N_R(A_j); 1 \leq j \leq m)$.
- Montrer que lorsque $R \rightarrow \infty$ la loi du processus ponctuel N_R converge vers celle d'un processus de Poisson ponctuel homogène d'intensité λ .

Exercice 5.2.2 Fixons un processus de Poisson ponctuel standard (c'est-à-dire homogène d'intensité 1) N sur \mathbb{R}^2 , et une fonction borélienne positive bornée f définie sur \mathbb{R} . Projétons sur l'axe horizontal de \mathbb{R}^2 les points du nuage \mathcal{E} définissant N qui sont situés entre l'axe horizontal et la courbe d'équation $y = f(x)$. Autrement dit, considérons l'ensemble aléatoire \mathcal{E}' des points $x \in \mathbb{R}$ pour lesquels existe $y \in [0, f(x)]$ tel que $(x, y) \in \mathcal{E}$. Soit N' le processus ponctuel associé à \mathcal{E}' .

- a) Vérifier que presque sûrement 2 points de \mathcal{E} ne peuvent se projeter sur le même point de \mathcal{E}' . (On pourra subdiviser tout carré $[-m, m]^2$ en bandes verticales de largeur $1/n$.)
- b) Montrer que N' est un processus de Poisson ponctuel, et donner son intensité.

Exercice 5.2.3 Fixons un processus de Poisson ponctuel sur \mathbb{R}_+ d'intensité μ admettant la densité continue h . Soit g une fonction borélienne localement bornée de \mathbb{R}_+ dans \mathbb{R} .

- a) Montrer que la loi du n -ième point S_n admet une densité, à exprimer (sans signe Σ).
- b) Montrer que $\mathbb{E}\left(\sum_{n \in \mathbb{N}^*} g(S_n) 1_{\{S_n \leq t\}}\right) = \int_0^t g h$, pour tout $t > 0$.

Exercice 5.2.4 (Processus de naissance et de mort) Soient N un processus de Poisson homogène réel d'intensité λ et $\{X_n; n \in \mathbb{N}^*\}$ une suite i.i.d. indépendante de N . Voyons S_n comme l'instant de naissance d'un n -ième individu, et X_n comme sa durée de vie. Notons $u \mapsto r(u) := \mathbb{P}(X_1 > u)$ la queue de la loi de la durée de vie, et Q_t le nombre des individus en vie à l'instant t .

- a) Montrer que $\mathbb{E}(a^{Q_t}) = \exp\left(-\lambda(1-a) \int_0^t r\right)$, pour $0 < a \leq 1$.
- b) Montrer que la loi de Q_t est $\mathcal{P}(\lambda \int_0^t r)$.
- c) Montrer que Q est un processus de Poisson ponctuel, et donner son intensité.

6 Files d'attente

6.1 Files à un serveur

Soient $\{X_n | n \in \mathbb{N}^*\}$ et $\{S_n | n \in \mathbb{N}^*\}$ deux suites indépendantes de v.a.i.i.d. à valeurs dans \mathbb{N}^* , de fonctions de répartition respectives F_X et F_S , et soit $T_n := X_1 + \dots + X_n$, pour $n \in \mathbb{N}^*$. (Et bien sûr $T_0 := 0$).

T_n est le temps d'arrivée dans la file du n -ième client, et S_n est le temps de service qu'il nécessite. Le client en tête de la file (i.e. le premier arrivé parmi ceux qui ne sont pas encore servis) sera systématiquement le premier servi (bien que d'autres possibilités soient envisageables a priori). Chaque client quitte la file dès qu'il a été servi.

La file ainsi définie est notée : $F_X/F_S/1$. (Le 1 signifiant qu'il n'y a qu'un serveur).

Les quantités qu'on étudie en priorité sont le nombre Q_t de clients présents dans la file (le client dont le service est en cours inclus) au temps $t \in \mathbb{N}$ (on part de $Q_0 = 0$, et on décompte immédiatement après t , de sorte que Q_t est continu à droite si t est vu comme variant dans \mathbb{R}_+), et le temps W_n que doit attendre le n -ième client avant de commencer à être servi.

Exemples 1) $\mathcal{G}_p/\mathcal{G}_q/1$, où \mathcal{G}_p désigne la loi géométrique de paramètre p . C'est la plus simple des files d'attente.

2) $\mathcal{G}_p/D(k)/1$, où $D(k)$ désigne une durée de service constante (déterministe) égale à $k \in \mathbb{N}^*$.

Notations Notons $D_0 := 0$, puis D_n le temps où le n -ième client quitte la file. Notons ensuite N_n le nombre de clients arrivant dans la file entre les instants 1 et n (inclus), et τ_n le nombre de clients quittant la file entre les instants 1 et n (inclus).

Notons encore $U_n := S_n - X_{n+1}$, pour tout $n \in \mathbb{N}^*$. C'est une suite de v.a.i.i.d. à valeurs dans \mathbb{Z} . Notons enfin F_n la fonction de répartition de la variable W_n .

Exercice n° 6.1.1 Vérifier les relations suivantes :

- $W_n = (W_{n-1} + S_{n-1} - X_n)^+$; $D_n = T_n + W_n + S_n = (D_{n-1} \vee T_n) + S_n$;
- $N_n = \max\{j \in \mathbb{N} \mid T_j \leq n\} = \sum_{j \geq 1} 1_{\{T_j \leq n\}}$;
- $\tau_n = \max\{j \in \mathbb{N} \mid D_j \leq n\} = \sum_{j \geq 1} 1_{\{D_j \leq n\}}$;
- $D_{\tau_n} \leq n < D_{1+\tau_n}$; $Q_n = N_n - \tau_n$; $Q_n > 0 \Leftrightarrow T_{1+\tau_n} \leq n$.

Exercice n° 6.1.2 a) Dédurre de l'exercice 6.1.1,a que pour tous $n-1, k \in \mathbb{N}$:

$$F_{n+1}(k) = \sum_{\{z \leq k\}} \mathbb{P}(U_1 = z) F_n(k-z) = \sum_{z \in \mathbb{N}} \mathbb{P}(U_1 = k-z) F_n(z).$$

b) Calculer F_1 et F_2 , et montrer que la suite (F_n) est décroissante.

c) Dédurre que la suite (F_n) converge vers une fonction F croissante, nulle sur \mathbb{R}_-^* , majorée par 1, et vérifiant $F(k) = \sum_{z \in \mathbb{N}} \mathbb{P}(U_1 = k-z) F(z)$ pour tout $k \in \mathbb{N}$.

d) Montrer que pour tous $n-2, k \in \mathbb{N}$:

$$F_n(k) = \mathbb{P}(U_1 \leq k, U_1 + U_2 \leq k, \dots, U_1 + \dots + U_{n-1} \leq k).$$

e) Dédurre que $F(k) = \mathbb{P}((\forall j \in \mathbb{N}^*) \sum_{i=1}^j U_i \leq k)$ pour tout $k \in \mathbb{N}$.

Proposition 22 *Supposons U_1 intégrable. Si $\mathbb{E}(U_1) > 0$, alors la suite (W_n) tend p.s. vers $+\infty$ (et la fonction F de l'exercice 6.1.2 ci-dessus est nulle). Tandis que si $\mathbb{E}(U_1) < 0$, alors la suite (W_n) converge en loi (et la fonction F de l'exercice 6.1.2 ci-dessus est la fonction de répartition de la limite). Enfin si $\mathbb{E}(U_1) = 0$, alors la suite (W_n) tend en probabilité vers $+\infty$ (et la fonction F de l'exercice 6.1.2 ci-dessus est nulle).*

Preuve Supposons d'abord $\mathbb{E}(U_1) > 0$. L'exercice 6.1.1,a et la loi des grands nombres montrent que p.s. $W_n/n \geq (U_1 + \dots + U_{n-1})/n \rightarrow \mathbb{E}(U_1)$, d'où $W_n \rightarrow +\infty$ (et $F_n \rightarrow 0$). Supposons ensuite $\mathbb{E}(U_1) < 0$, et utilisons l'exercice 6.1.2 et la loi des grands nombres.

Fixons $\delta > 0$ et $\varepsilon \in]0, -\mathbb{E}(U_1)[$, puis $N = N(\varepsilon, \delta)$ tel que

$$1 - \delta < \mathbb{P}((\forall n \geq N) \mid \sum_{j=1}^n U_j/n - \mathbb{E}(U_1) \mid < \varepsilon) \leq \mathbb{P}((\forall n \geq N) \sum_{j=1}^n U_j \leq 0).$$

Notons $A_k := \cap_{n=1}^{N-1} \{\sum_{j=1}^n U_j \leq k\}$, et $B := \cap_{n \geq N} \{\sum_{j=1}^n U_j \leq 0\}$, puis fixons k suffisamment grand pour que $\mathbb{P}(A_k) > 1 - \delta$. Alors

$\lim_{\infty} F \geq F(k) = \mathbb{P}(\cap_{n \geq 1} \{\sum_{j=1}^n U_j \leq k\}) \geq \mathbb{P}(A_k \cap B) \geq 2(1 - \delta) - 1 = 1 - 2\delta$, et comme δ est quelconque > 0 , on a $\lim_{\infty} F = 1$; par ailleurs, comme limite décroissante d'une suite de fonctions croissantes, F est càd ; ce qui assure que F est une fonction de répartition, disons d'une certaine v.a. V . Par conséquent W_n converge en loi vers V .

Supposons enfin que $\mathbb{E}(U_1) = 0$. Alors d'après le théorème 3 (de Chung et Fuchs, section 2.4) la marche aléatoire centrée $\sum_{j=1}^n U_j$ est récurrente (sur le sous-groupe engendré par la réunion des supports des lois de X_1 et de S_1 , voir la proposition 3 et l'exercice 2.2.2). En particulier, $\max\{\sum_{j=1}^n U_j \mid n \leq \ell\}$ tend p.s. vers $+\infty$ avec ℓ , et donc F_n (voir l'exercice 6.1.2,d) tend vers 0, ce qui assure la convergence en probabilité vers ∞ par définition de F_n . \diamond

6.2 Files $\mathcal{G}_p/\mathcal{G}_q/1$

Exercice n° 6.2.1 Supposons que X_1 est géométrique de paramètre $p \in]0, 1[$.

a) Montrer que pour $k \in \mathbb{N}^*$ et $1 \leq n_1 < n_2 < \dots < n_k$, on a

$$\mathbb{P}(T_1 = n_1, T_2 = n_2, \dots, T_k = n_k) = p^k (1-p)^{n_k-k}.$$

b) Montrer que pour $k \in \mathbb{N}^*$, $1 \leq b_1 < \dots < b_k$ et $1 \leq n_{b_1} < \dots < n_{b_k}$, on a

$$\mathbb{P}(T_{b_1} = n_{b_1}, \dots, T_{b_k} = n_{b_k}) = p^{b_k} (1-p)^{n_{b_k}-b_k} \times C_{n_{b_1}-1}^{b_1-1} \times \dots \times C_{n_{b_k}-1-n_{b_{k-1}}}^{b_k-1-b_{k-1}};$$

vérifier que c'est encore vrai pour $0 \leq b_1 \leq \dots \leq b_k$ et $0 \leq n_{b_1} \leq \dots \leq n_{b_k}$, en posant $C_{a-1}^{-1} := 1_{\{a=0\}}$ et $C_{a-1}^{b-1} = 0$ pour tout $a \in \mathbb{N}$ et $b > a$.

c) Montrer que pour $a_1 \in \mathbb{N}$, $k, a_2, \dots, a_k \in \mathbb{N}^*$ et $0 \leq n_1 \leq n'_1 \leq n_2 \leq \dots \leq n_k \leq n'_k$:

$$\begin{aligned} \mathbb{P}(T_{a_1} = n_1, T_{a_1+a_2} = n'_1, T_{a_1+a_2} = n_2, \dots, T_{a_1+\dots+a_k} = n_k, T_{a_1+\dots+a_k+1} = n'_k) \\ = \left(\frac{p}{1-p}\right)^{a_1+\dots+a_k+1} (1-p)^{n'_k} \times C_{n_1-1}^{a_1-1} C_{n_2-n'_1-1}^{a_2-2} \dots C_{n_k-n'_{k-1}-1}^{a_k-2}. \end{aligned}$$

d) Montrer que $\sum_{n=0}^m C_{n-1}^{a-1} = C_m^a$, pour m et a dans \mathbb{N} , puis que

$$\sum_{m' < n' \leq n \leq m} C_{n-n'-1}^{a-2} = C_{m-m'}^a, \text{ pour } a \in \mathbb{N}^* \text{ et } 0 \leq m' \leq m.$$

e) Dédire que pour $a_1 \in \mathbb{N}$, $k, a_2, \dots, a_k \in \mathbb{N}^*$ et $0 \leq m_1 \leq m'_1 \leq m_2 \leq \dots \leq m_k \leq m'_k$:

$$\begin{aligned} \mathbb{P}(T_{a_1} \leq m_1, m'_1 < T_{a_1+1}, \dots, T_{a_1+\dots+a_k} \leq m_k, m'_k < T_{a_1+\dots+a_k+1}) \\ = \left(\frac{p}{1-p}\right)^{a_1+\dots+a_k} (1-p)^{m'_k} \times C_{m_1}^{a_1} C_{m_2-m'_1}^{a_2} \dots C_{m_k-m'_{k-1}}^{a_k}. \end{aligned}$$

f) Dédire que pour $k \in \mathbb{N}^*$, $\alpha_1, \dots, \alpha_k \in \mathbb{N}$, et $0 \leq n_1 \leq n_2 \leq \dots \leq n_k$ on a :

$$\mathbb{P}(N_{n_1} = \alpha_1, \dots, N_{n_k} - N_{n_{k-1}} = \alpha_k) = \left(\frac{p}{1-p}\right)^{\alpha_1+\dots+\alpha_k} (1-p)^{n_k} \times C_{n_1}^{\alpha_1} C_{n_2-n_1}^{\alpha_2} \dots C_{n_k-n_{k-1}}^{\alpha_k}.$$

g) Dédire que pour $k \in \mathbb{N}^*$, $\alpha_1, \dots, \alpha_k \in \mathbb{N}$ et $0 \leq n_0 \leq n_1 \leq \dots \leq n_k$ on a :

$$\mathbb{P}(N_{n_1} - N_{n_0} = \alpha_1, \dots, N_{n_k} - N_{n_{k-1}} = \alpha_k) = \left(\frac{p}{1-p}\right)^{\alpha_1+\dots+\alpha_k} (1-p)^{n_k-n_0} C_{n_1-n_0}^{\alpha_1} \dots C_{n_k-n_{k-1}}^{\alpha_k}.$$

h) Dédire que $\{N_n \mid n \in \mathbb{N}\}$ est une marche aléatoire sur \mathbb{N} , issue de 0, dont on précisera la loi de déplacement.

Exercice n° 6.2.2 Supposons que X_1 est géométrique de paramètre $p \in]0, 1[$.

a) Montrer que $\min\{n \in \mathbb{N} \mid N_n \geq k\} = T_k$, pour tout $k \in \mathbb{N}$.

b) Montrer que pour $j \in \mathbb{N}$ et $n > n_1 > n_2 > \dots > n_{j+1} = 1$:

$$\mathbb{P}(N_n = N_{n-1} = \dots = N_{n_1} = j, N_{n_1-1} = \dots = N_{n_2} = j-1, \dots, N_{n_j-1} = \dots = N_{n_{j+1}} = 0)$$

$$= \mathbb{P}(T_1 = n_j, T_2 = n_{j-1}, \dots, T_j = n_1, T_{j+1} > n)$$

$$= (1-p)\mathbb{P}(N_{n-1} = \dots = N_{n_1} = j, N_{n_1-1} = \dots = N_{n_2} = j-1, \dots, N_{n_{j-1}-1} = \dots = N_{n_{j+1}} = 0).$$

c) D duire que pour $n \in \mathbb{N}^*$ nous avons

$$\mathbb{P}(N_n - N_{n-1} = 0 \mid N_{n-1}, N_{n-2}, \dots, N_1) = \mathbb{P}(N_n - N_{n-1} = 0) = 1 - p,$$

$$\text{et } \mathbb{P}(N_n - N_{n-1} = 1 \mid N_{n-1}, N_{n-2}, \dots, N_1) = \mathbb{P}(N_n - N_{n-1} = 1) = p.$$

d) D duire de l  le r sultat de l'exercice 6.1.3 (h). Comparer les deux m thodes.

En g n ral, Q_n n'est pas une cha ne de Markov. C'est cependant le cas pour les files $\mathcal{G}_p/\mathcal{G}_q/1$, du fait de la propri t  d'absence de m moire des variables g om triques : nous avons en effet pour tout $\ell \in \mathbb{N}$: $\mathbb{P}(X_1 = \ell) = p \mathbb{P}(X_1 \geq \ell)$ et $\mathbb{P}(S_1 = \ell) = q \mathbb{P}(S_1 \geq \ell)$.

Proposition 23 *Dans le cas de la file $\mathcal{G}_p/\mathcal{G}_q/1$, la suite Q_n d finie par la longueur de la file   l'instant n forme une cha ne de Markov sur \mathbb{N} , dont la matrice de transition est d finie par : $P(0, 1) = p = 1 - P(0, 0)$ et pour tout $k \geq 1$: $P(k, k - 1) = q(1 - p)$, $P(k, k + 1) = p(1 - q)$, $P(k, k) = pq + (1 - p)(1 - q)$.*

Preuve Fixons $n \in \mathbb{N}^*$, et j, k dans \mathbb{N} . Nous avons

$$\mathbb{P}(\tau_n > \tau_{n-1}; \tau_{n-1} = \dots = \tau_{n_1} = j; \tau_{n_1-1} = \dots; N_n = N_{n-1} = \dots = N_{m_1} = k; N_{m_1-1} = \dots)$$

$$= \mathbb{P}(D_{j+1} = n; D_j = n_1; \dots; T_{k+1} > n; T_k = m_1; \dots; T_{j+1} = m_{k-j}; \dots) 1_{\{j < k\}}$$

(car $D_m > T_m$ pour tout m , ce qui impose $k > j$ ci-dessus ; l'indicatrice de $\{j < k\}$ est sous-entendue ci-dessous)

$$= \mathbb{P}(S_{j+1} = n - n_1 \vee m_{k-j}; D_j = n_1; \dots; X_{k+1} > n - m_1; T_k = m_1; \dots; T_{j+1} = m_{k-j}; \dots)$$

$$= \mathbb{P}(S_{j+1} = n - n_1 \vee m_{k-j}) \mathbb{P}(X_{k+1} > n - m_1) \mathbb{P}(D_j = n_1; \dots; T_k = m_1; \dots; T_{j+1} = m_{k-j}; \dots)$$

$$= q \mathbb{P}(S_{j+1} \geq n - n_1 \vee m_{k-j}) (1-p) \mathbb{P}(X_{k+1} \geq n - m_1) \mathbb{P}(D_j = n_1; \dots; T_k = m_1; \dots; T_{j+1} = m_{k-j}; \dots)$$

$$= q(1-p) \mathbb{P}(S_{j+1} \geq n - n_1 \vee m_{k-j}; D_j = n_1; \dots; X_{k+1} \geq n - m_1; T_k = m_1; \dots; T_{j+1} = m_{k-j}; \dots)$$

$$= q(1-p) \mathbb{P}(\tau_{n-1} = \dots = \tau_{n_1} = j; \tau_{n_1-1} = \dots; N_{n-1} = \dots = N_{m_1} = k; N_{m_1-1} = \dots) 1_{\{j < k\}}.$$

En sommant sur toutes les possibilit s relatives   $N_{n-1}, N_{n-2}, \dots, \tau_{n-1}, \tau_{n-2}, \dots$, on d duit :

$$\mathbb{P}(\tau_n > \tau_{n-1}; N_n = N_{n-1}; Q_{n-1} = k - j; Q_{n-2} = \dots) = q(1-p) \mathbb{P}(Q_{n-1} = k - j; Q_{n-2}; \dots) 1_{\{j < k\}},$$

c'est- -dire

$$\mathbb{P}(Q_n = Q_{n-1} - 1 \mid Q_{n-1}; Q_{n-2}; \dots) = q(1-p) 1_{\{Q_{n-1} > 0\}}.$$

De la m me fa on, on obtient

$$\mathbb{P}(Q_n = Q_{n-1} + 1 \mid Q_{n-1}; Q_{n-2}; \dots) = p(1-q) 1_{\{Q_{n-1} > 0\}} + p 1_{\{Q_{n-1} = 0\}}.$$

Enfin, comme $|Q_n - Q_{n-1}| \leq 1$, nous avons aussi

$$\mathbb{P}(Q_n = Q_{n-1} \mid Q_{n-1}; Q_{n-2}; \dots) = 1 - q(1-p) 1_{\{Q_{n-1} > 0\}} - p(1-q) 1_{\{Q_{n-1} > 0\}} - p 1_{\{Q_{n-1} = 0\}} \cdot \diamond$$

Exercice n° 6.2.3 Reprenant la preuve ci-dessus, démontrer en détail que

$$\mathbb{P}(Q_n = Q_{n-1} + 1 \mid Q_{n-1}; Q_{n-2}; \dots) = p(1-q) 1_{\{Q_{n-1} > 0\}} + p 1_{\{Q_{n-1} = 0\}} \cdot$$

Exercice n° 6.2.4 Toujours dans le cas $\mathcal{G}_p/\mathcal{G}_q/1$:

- Montrer que les mesures invariantes sont proportionnelles à ν définie par $\nu_k = \left(\frac{p(1-q)}{q(1-p)}\right)^k$.
- Dire quand la chaîne Q_n est positivement récurrente.
- Comparant $\mathbb{P}_0(T_0 = \infty)$ avec $\mathbb{P}_1(T_0 = \infty)$, puis la chaîne Q_n avec une marche aléatoire, déduire du théorème 3 (de Chung et Fuchs, section 2.4) quand la chaîne Q_n est récurrente.
- Lorsque $q > p$, montrer que sous la probabilité invariante la loi de W_n est géométrique, de paramètre constant à calculer.

6.3 Files $\mathcal{G}_p/F_S/1$

Pour ces files, la suite Q_n n'est markovienne que si S_1 est géométrique, comme dans le cas précédemment étudié. Mais on a encore droit aux résultats des exercices 6.2.1 et 6.2.2, décrivant le processus N_n comme une marche aléatoire sur \mathbb{N} . Et cela entraîne le résultat suivant, qui énonce essentiellement que, à défaut d'être markovienne, la suite Q_n contient une sous-suite markovienne.

Proposition 24 *Supposons X_1 géométrique de paramètre $p \in]0, 1[$. Alors la suite Q_{D_n} est markovienne sur \mathbb{N} . Sa matrice de transition est donnée par :*

$$P(j, k) = \mathbb{P}(N_{S_1} = k - (j-1)^+) = \left(\frac{p}{1-p}\right)^{k-(j-1)^+} \times \mathbb{E}\left((1-p)^{S_1} C_{S_1}^{k-(j-1)^+}\right).$$

Preuve Si $Q_{D_n} > 0$, alors $Q_{D_{n+1}} = Q_{D_n} + (N_{D_n+S_{n+1}} - N_{D_n}) - 1$; et si $Q_{D_n} = 0$, alors $Q_{D_{n+1}} = N_{T_{n+1}+S_{n+1}} - N_{T_{n+1}}$. Donc utilisant que N_n est une marche aléatoire (voir l'exercice 6.2.2), on déduit que $Q_{D_{n+1}} - Q_{D_n}$ a la loi de $N_{S_1} - 1_{\{Q_{D_n} > 0\}}$.

Ceci prouve bien que Q_{D_n} est markovienne, et énonce de plus que pour tout $k \in \mathbb{N}$:

$$P(j, k) = \mathbb{P}(N_{S_1} - 1 = k - j) = \mathbb{P}(N_{S_1} = k - (j-1)^+) \text{ si } j \in \mathbb{N}^*, \text{ et}$$

$$P(0, k) = \mathbb{P}(N_{S_1} = k) = \mathbb{P}(N_{S_1} = k - (0-1)^+).$$

Enfin la loi explicite de N_{S_1} découle aussitôt de l'exercice 6.2.1.f. \diamond

Proposition 25 *Supposons X_1 géométrique de paramètre $p \in]0, 1[$, et notons $\rho := p\mathbb{E}(S_1)$ la "densité de trafic". Alors la chaîne de Markov Q_{D_n} est transitoire lorsque $\rho > 1$, nulle-récurrente lorsque $\rho = 1$, et ergodique lorsque $\rho < 1$.*

Preuve La proposition 24 ci-dessus montre que Q_{D_n} est irréductible et apériodique, et qu'on peut écrire $Q_{D_n} = Y_1 + \dots + Y_n - n + Z_n$, où les Y_j sont i.i.d. de loi égale à celle de N_{S_1} , et où $Z_n := \sum_{j=0}^{n-1} 1_{\{Q_{D_j}=0\}}$. Par ailleurs comme de la loi de N_n on tire aisément que son espérance est np , on voit que $\varrho = \mathbb{E}(N_{S_1})$, et donc que $0 \leq \mathbb{E}(Q_{D_n}) = n(\varrho - 1) + \mathbb{E}(Z_n)$. Alors utilisant le corollaire 6 de la section 3.5 nous déduisons : $(1 - \varrho) \leq \mathbb{E}(Z_n/n) \rightarrow 1/m_0$, où m_0 est l'espérance du temps de retour en 0 pour la chaîne Q_{D_n} (issue de 0).

Par conséquent, $\varrho < 1$ entraîne $m_0 < \infty$, et donc l'ergodicité.

D'autre part, la loi des grands nombres et la proposition 17 de la section 3.5 montrent que p.s. $Q_{D_n}/n \rightarrow \varrho - 1 + 1/m_0 \geq \varrho - 1$, et donc que $Q_{D_n} \rightarrow +\infty$ p.s. si $\varrho > 1$, d'où la transience. Supposons enfin $\varrho = 1$. Alors le théorème 3 (de Chung et Fuchs, section 2.4) assure que $Q_{D_n} - Z_n$ est une marche récurrente, supportée par un sous-groupe $\neq \{0\}$ de \mathbb{Z} (voir la proposition 3 de la section 2.2). En particulier, nous avons donc p.s. $\liminf_n (Q_{D_n} - Z_n) = -\infty$, ce qui exclut que Z_n soit borné, et prouve donc la récurrence. Pour conclure, si la récurrence n'était pas nulle alors m_0 serait fini, et nous aurions p.s. $Q_{D_n}/n \rightarrow 1/m_0 > 0$, d'où $Q_{D_n} \rightarrow +\infty$, ce qui contredirait la récurrence. \diamond

Exercice n° 6.3.1 Notons G la fonction génératrice de S_1 , et A celle de N_{S_1} .

a) Montrer que $A(s) = G(1 - p + ps)$, pour tout $s \in [0, 1]$, et que $A'(1-) = \varrho$.

Supposons dorénavant que $\varrho < 1$, et notons ν la probabilité invariante de Q_{D_n} .

b) Soit $H(s) := \mathbb{E}_\nu(s^{Q_{D_n}}) = \sum_{j \in \mathbb{N}} \nu_j s^j$. Montrer (en utilisant la preuve de la proposition 25 ci-dessus) que $sH(s) = A(s) \times (\nu_0 s + \sum_{n \in \mathbb{N}^*} \nu_n s^n)$, puis déduire que $\nu_0 = 1 - \varrho$ et que $H(s) = (1 - \varrho)(1 - s)A(s)/(A(s) - s)$.

c) En calculant $H'(1-)$ à partir de b), montrer que $\mathbb{E}_\nu(Q_{D_n}) = \varrho + p^2 \mathbb{E}(S_1^2)/(2(1 - \varrho))$.

d) Vérifier que $Q_{D_n} = N_{D_n} - N_{T_n}$, en déduire que

$\mathbb{E}(s^{Q_{D_n}}) = \mathbb{E}((1 - p + ps)^{W_n}) \times \mathbb{E}((1 - p + ps)^{S_n})$, et puis que

$\mathbb{E}_\nu(s^{W_n}) = (1 - \varrho)(1 - s) \times [1 - s - p(1 - G(s))]^{-1}$ et $\mathbb{E}_\nu(W_n) = p \mathbb{E}(S_1^2)/(2(1 - \varrho))$.

Bibliographie

Plus ou moins aisément lisible

BAUER H. *Probability theory.* de Gruyter, Berlin-New-York 1996.

BRÉMAUD P. *Introduction aux probabilités.* Springer, Berlin 1984.

CHUNG K.L. *A course in probability theory.* Academic Press 1974.

COTTRELL M., GENON-CATALOT V., DUHAMEL C., MEYRE T.

Exercices de probabilités. Cassini, Paris 1999.

DEL MORAL P., RÉMILLARD B., RUBENTHALER S.

Une introduction aux probabilités. Ellipses 2006.

- FOATA D., FUCHS A. *Calcul des probabilités.* Dunod, Paris 1998.
- FOATA D., FUCHS A. *Processus stochastiques.* Dunod, Paris 2002.
- LETAC G. *Intégration et probabilités.* Masson, Paris 1997.
- MAZLIAK M., PRIOURET P., BALDI P. *Chaînes de Markov, exercices.* Hermann 1998.
- MÉTIVIER M., NEVEU J. *Cours de probabilités.* École polytechnique, Palaiseau 1979.
- PARDOUX É. *Processus de Markov et applications.* Dunod, Paris 2007.
- REVUZ D. *Probabilités.* Hermann, Paris 1994.
- ROSS S. *Introduction to probability models.* Academic press 1980.
- VALLOIS P. *Modélisations stochastiques et simulations.* Ellipses, Paris 2007.
- Plus difficilement lisible
- BACCELLI F., BRÉMAUD P. *Elements of queueing theory.* Springer, Berlin 1994.
- BENAÏM M., EL KAROUI N. *Promenade aléatoire.*
Éditions de l'école polytechnique, Palaiseau 2004.
- BOULEAU N. *Processus stochastiques et applications.* Hermann, Paris 1988.
- BRÉMAUD P. *Point processes and queues.* Springer, Berlin 1981.
- CHUNG K.L. *Markov chains with stationary transition probabilities.* Springer, 1960.
- DURRETT R. *Probability : theory and examples.* Duxbury Press 1996.
- FELLER W. *An introduction to probability theory and its applications.*
Vol. 1, Wiley and sons, New-York 1968.
- GRAHAM C. *Chaînes de Markov.* Dunod, Paris 2008.
- GRIMMET G., STIRZAKER D. *Probability and random processes.*
Clarendon press, Oxford 1992.
- KARLIN S. *Initiation aux processus aléatoires.* Dunod, Paris 1969.
- KARLIN S., TAYLOR H. *A first course in stochastic processes.*
Academic Press, New-York 1975.
- LAWLER G.F. *Introduction to stochastic processes.* Chapman and Hall, Londres 1995.
- NEVEU J. *Martingales à temps discret.* Masson, Paris 1972.

- NORRIS J.R. *Markov chains.* Cambridge university press, Cambridge 1997.
- ROBERT P. *Réseaux et files d'attente : méthodes probabilistes.* Springer, Berlin 2000.
- SENETA E. *Non-negative matrices and Markov chains.* Springer Verlag, Berlin 1973.
- TOULOUSE P.S. *Thèmes de probabilités et statistiques.* Dunod, Paris 1999.
- WILLIAMS D. *Probabilities with martingales.* Cambridge university press, 1991.
- WOESS W. *Catene di Markov e teoria del potenziale nel discreto.*
Pitagora editrice, Bologna 1996.