Philippe CARMONA

NOTES DU COURS PROCESSUS STOCHASTIQUES

Laboratoire Jean Leray, UMR 6629, Université de Nantes, 92208, F-44322, Nantes cedex 03,
e-mail: philippe.carmona@math.univ-nantes.fr

Philippe Carmona

Laboratoire Jean Leray, UMR 6629, Université de Nantes, 92208, F-44322, Nantes cedex 03,
e-mail: philippe.carmona@math.univ-nantes.fr

NOTES DU COURS PROCESSUS STOCHASTIQUES

Philippe Carmona

Résumé. — Ces notes sont un support pour le cours d'introduction aux processus stochastiques, de l'option mathématique de l'ECN Nantes. Nous conseillons la lecture d'ouvrages plus conséquents:

- "Théorèmes limites et processus de Poisson", Grégory Miermont, `http://perso.ens-lyon.fr/gregory.miermont/thlim.pdf`
- "Essentials of stochastic processes", R. Durrett , Springer
- "A first course in stochastic models", H. Tijms, Wiley.

Laboratoire Jean Leray, UMR 6629, Université de Nantes, 92208, F-44322, Nantes cedex 03, e-mail: philippe.carmona@math.univ-nantes.fr

1 Processus de Poisson

SECTION 1

La loi exponentielle

Définition 1.1. — *On dit que la variable aléatoire réelle X suit une loi exponentielle de paramètre $\lambda > 0$, et on note $X \sim exp(\lambda)$ si elle admet pour densité $f_\lambda(x) = \lambda e^{-\lambda x} \mathbf{1}_{(x>0)}$.*

On vérifie sans peine que la fonction de répartition de X est

$$F(t) = \mathbb{P}(X \leq t) = 1 - e^{-\lambda t} \quad (t \geq 0),$$

que X est de carré intégrable, $\mathbb{E}[X] = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$ et que $\frac{X}{\lambda} \sim exp(1)$.

***Proposition 1.1* (Manque de mémoire de la loi exponentielle)**
Soit X une variable aléatoire réelle telle que $0 < X < +\infty$ p.s. Alors X satisfait l'équation fonctionnelle

$$\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s) \quad (s, t \geq 0)$$

ssi il existe $\lambda > 0$ tel que $X \sim exp(\lambda)$.

Démonstration. — Il est trivial de vérifier que la loi exponentielle vérifie l'équation fonctionnelle car $\mathbb{P}(X > s) = e^{-\lambda s}$, pour $s \geq 0$.
Réciproquement si X satisfait l'équation fonctionnelle $f(t+s) = f(s)f(t)$ avec $f(s) = \mathbb{P}(X > s)$, alors on a $f(s) > 0$ pour tout $s \geq 0$ car si $f(s) = 0$, alors $f(s/2) = \sqrt{f(s)} = 0$ et donc $1 = f(0) = \lim_{n\to+\infty} f(s/2^n) = 0$ ce qui est contradictoire. On peut donc considérer la fonction $g(t) = lnf(t)$ qui vérifie l'équation de Cauchy

$$g(t + s) = g(t) + g(s) \quad (s, t \geq 0).$$

Il est bien connu que comme g est localement intégrable, il existe une constante réelle a telle que $g(x) = ax$, ce qui donne $f(t) = e^{at}$. Comme $X < +\infty$ ps, il existe t tel que $f(t) < 1$ ce qui montre que $a < 0$. $\square$

***Proposition 1.2* (La course d'exponentielles)**. — *Soit X, Y deux variables aléatoires indépendantes, $X \sim exp(\lambda)$, $Y \sim \exp(\mu)$. Alors $Z = \inf(X, Y) \sim exp(\lambda + \mu)$.*

Démonstration. — Pour $t \geq 0$, $\mathbb{P}(Z > t) = \mathbb{P}(X > t, Y > t) = e^{-\lambda t} e^{-\mu t}$. $\square$

Si on fait une course, on désire non seulement connaître le temps du gagnant, mais aussi l'identité du gagnant. Soit donc $X_1, \ldots, X_n$ des variables aléatoires indépendantes qui suivent des lois exponentielles $X_i \sim exp(\lambda_i)$. Alors $Z = \inf(X_1, \ldots, X_n)$ est le temps du gagnant. Observons que par indépendance, si $i \neq j$, alors

$$\mathbb{P}(X_i = X_j) = \int \mathbf{1}_{(x=y)} f_{\lambda_i}(x) f_{\lambda_j}(y) dx\, dy = 0$$

et donc, par additivité,

$$\mathbb{P}(\exists i \neq j, X_i = X_j) \leq \sum_{i \neq j} \mathbb{P}(X_i = X_j) = 0\,.$$

On se placera donc sur un ensemble de probabilité 1,

$$\Omega' = \{\omega : \forall i \neq j, X_i(\omega) \neq X_j(\omega)\}$$

sur lequel la variable aléatoire numéro du gagnant est bien définie:

$$N = \inf\{i \leq n : X_i = Z\}.$$

***Proposition 1.3* (Course d'exponentielles générale)**
Les variables aléatoires N et Z sont indépendantes,

$$Z \sim exp(\bar{\lambda}) \quad et \quad \mathbb{P}(N = i) = \frac{\lambda_i}{\bar{\lambda}} \quad (1 \leq i \leq n)\,.$$

Démonstration. — On établit que pour $t \geq 0$ et $1 \leq i \leq n$, on a

$$\mathbb{P}(Z \geq t, N = i) = e^{-\bar{\lambda} t} \frac{\lambda_i}{\bar{\lambda}}\,.$$

On fait $t = 0$ pour obtenir la loi de N, puis on en déduit que, sous la loi conditionnelle $\mathbb{P}(. \mid N = i)$ la fonction de répartition de Z est celle de la loi $\nu = \exp(\bar{\lambda})$. D'ou le fait que la loi conditionnelle de Z sachant $N = i$ est ν. On conclut en écrivant que pour f, g mesurables positives:

$$\mathbb{E}[f(Z)h(N)] = \sum_i \mathbb{E}[f(Z) \mid N = i] h(i) \mathbb{P}(N = i) = \sum_i \nu(f) h(i) \mathbb{P}(N = i) = \nu(f) \mathbb{E}[h(N)]\,.$$

$\square$

***Exemple 1.1.* — Une autre histoire de bus.** *Pour se rendre de la gare à l'aéroport on a le choix entre deux lignes, la 1 et la 2. Les bus de la ligne 1 arrivent à l'arrêt suivant un processus de Poisson de paramètre λ_1, ceux de la ligne 2 suivant un processus de Poisson de paramètre λ_2. Les deux processus sont supposés indépendants. Les temps de parcours entre la gare et l'aéroport sont supposés déterministes, et valent respectivement t_1 et t_2 minutes. On supposera pour les applications numériques que $\lambda_1 = 1/5$ et $\lambda_2 = 1/10$.*
John est un adepte de la ligne 1, car il trouve que les bus passent plus souvent : lorsqu'il arrive à la gare, il attend un bus de la ligne 1, puis se rend à l'aéroport. Sarah elle est une adepte de la ligne 2, car elle trouve que les bus vont plus vite : elle attend un bus de la ligne 2, puis se rend à l'aéroport.

1. *Déterminer les temps de transport moyen de John et Sarah. En déduire une Condition nécessaire et Suffisante, portant sur t_1, t_2 pour que John ait la meilleure stratégie : sur un grand nombre de voyages il met en moyenne moins de temps que Sarah. Tracer, dans le plan (t_1, t_2), la zone où John a raison.*
2. *Une troisième personne, Bill, a une stratégie différente : elle prend le premier bus disponible. Calculer son temps moyen de parcours, et tracer dans le plan les zones ou John, Bill et Sarah ont la meilleure stratégie.*
3. *Application numérique $t_1 = 25$. Indiquer suivant les valeurs de t_2, quelle est la meilleure stratégie. Interpréter vos résultats.*

Remarque. — *Cet exercice utilise en premier lieu la loi des grands nombres pour se ramener à comparer des moyennes. La réponse à la première question est élémentaire pour une personne douée de bon sens : si $t_1 = 25$, comme les bus de la ligne* 1 *passent en moyenne toutes les 5 minutes, le temps moyen de transport de John est de 30 mn. Si $t_2 > 20$, alors Sarah mettra en moyenne plus de 30 mn, et donc John a la meilleure stratégie. Il est plus difficile de comparer la stratégie de Bill aux autres car son temps moyen d'attente est* $3,33$ *mn, facile à calculer mais difficile à deviner!*

La loi Gamma On rappelle que la loi $\gamma(a, b)$ de paramètre d'échelle $a > 0$ et de paramètre de forme $b > 0$ admet pour densité:

$$f_{a,b}(x) = b^a \frac{x^{a-1}}{\Gamma(a)} e^{-bx} \mathbf{1}_{(x>0)}.$$

avec $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x}\, dx$ qui vérifie la relation $\Gamma(a+1) = a\Gamma(a)$ (on en déduit $\Gamma(n+1) = n!$).

***Lemme 1.4* (additivité des lois gamma).** —

$$\gamma(a, b) * \gamma(a', b) = \gamma(a + a', b).$$

Démonstration. — Il suffit de calculer la transformée de Laplace de la loi γ. Si $X \sim \gamma(a,b)$, alors

$$\mathbb{E}\left[e^{sX}\right] = \begin{cases} +\infty & si\, s \geq b \\ \left(\frac{b}{b-s}\right)^a & sinon. \end{cases}$$

□

Corollaire 1.5. — *Soit $X_1, \ldots, X_n$ indépendantes de même loi $exp(\lambda)$. Alors $S_n = X_1 = \cdots + X_n$ suit la loi $\gamma(n,\lambda)$ et donc admet pour densité $\lambda^n \frac{x^{n-1}}{(n-1)!} e^{-\lambda x}\, \mathbf{1}_{(x>0)}$.*

Démonstration. — On note que $exp(\lambda) = \gamma(\lambda, 1)$ et on applique le Lemme précédent. □

La loi de Poisson Rappelons qu'une variable aléatoire entière N suit la loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$, si

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

On montre aisément que $\mathbb{E}[N] = \mathrm{Var}(N) = \lambda$, par exemple en calculant la fonction génératrice, la fonction caractéristique ou la transformée de Laplace: pour $|u| \leq 1; s,t \in \mathbb{R}$, on a :

$$G_N(u) = \mathbb{E}\left[u^N\right] = e^{\lambda(u-1)}, \quad \mathbb{E}\left[e^{sN}\right] = e^{\lambda(e^s-1)} \quad \mathbb{E}\left[e^{itN}\right] = e^{\lambda(e^{it}-1)}.$$

On déduit aussi immédiatement de la forme de la fonction caractéristique l'additivité des lois de Poisson:

$$\mathcal{P}(\lambda) * \mathcal{P}(\mu) = \mathcal{P}(\lambda + \mu).$$

SECTION 2

Le Processus de Poisson homogène

Le processus de Poisson est un processus de comptage : on compte le nombre d'occurences au cours du temps d'un évènement spécifique (arrivée d'un appel téléphonique, entrée d'un client dans une boutique, etc ...). Il apparaît naturellement comme processus limite (voir section 4). Faire un dessin! Plus rigoureusement,

Définition 1.2. — *Un Processus de Poisson de paramètre λ est une famille de variables aléatoires $(N_t)_{t\geq 0}$ à valeurs dans $\mathbb{N}$ telle que*

(a) la fonction $t \to N_t$ est croissante, continue à droite et ne croît que par sauts de 1.

(b) Pour tous $s,t \geq 0$, $N_{t+s} - N_s \sim \mathcal{P}(\lambda t)$.

(c) *Si $t_0 < t_1 < \ldots < t_n$, alors les variables aléatoires $(N_{t_{i+1}} - N_{t_i})_{0\leq i\leq n-1}$ sont indépendantes.*

Remarque. — *Un processus qui vérifie (c) est dit à accroissements indépendants. La propriété (b) dit que les accroissements sont stationnaires (et poissoniens).*
N_t est donc le nombre d'évènements spécifiques qui se sont produits dans l'intervalle $[0, t]$.

On montrera plus tard comme conséquence du théorème de construction d'une mesure de Poisson générale, que si on se donne des variables aléatoires $(X_n)_{n\in\mathbb{N}}$ indépendantes de même loi $exp(\lambda)$ et si on pose

$$S_0 = 0, S_n = X_1 + \cdots + X_n, n \geq 1, \quad N_t = \sum_{n\geq 1} \mathbf{1}_{(S_n \leq t)}$$

alors $(N_t)_{t\geq 0}$ est un Processus de Poisson de paramètre λ. Les variables S_n sont appelées instants de saut du processus N, et on observe que par construction

$$\{N_t \geq n\} = \{S_n \leq t\} \qquad \text{et donc} \quad \{N_t = n\} = \{S_n \leq t < S_{n+1}\}.$$

Proposition 1.6. — *Conditonnellement à $N_t = n$, $(S_1, \ldots, S_n)$ a même loi que $(U'_1, \ldots, U'_n)$ statistiques s'ordre de $(U_1, \ldots, U_n)$ IID de loi uniforme sur $[0, t]$. Cette loi nommée $\mathcal{D}_{n,t}$ est appélée loi de Dirichlet $\mathcal{D}_{n,t}$ et admet pour densité*

$$g_{n,t}(s_1, \ldots, s_n) = \frac{n!}{t}\mathbf{1}_{(0<s_1<s_2<\ldots<s_n<t)}.$$

SECTION 3

Construction d'un processus de Poisson général

Définition 1.3. — *On appelle mesure ponctuelle sur un espace mesurable $(E, \mathcal{E})$ une somme finie ou dénombrable de masses de Dirac.*

On note $\mathcal{M}_p(E)$ l'ensemble des mesures ponctuelles sur $(E, \mathcal{E})$. On notera $\mu = \sum_{x\in D} \delta_x$ une mesure $\mu \in \mathcal{M}_p(E)$: D est un ensemble fini ou dénombrable. On dit que la mesure μ est *simple* si chaque point x apparait auplus une fois dans la somme.
On munit $\mathcal{M}_p(E)$ de la plus petite tribu rendant mesurables les fonctions $\mu \to \mu(f) = \sum_{x\in D} f(x)$ avec $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable positive.

Définition 1.4. — *Etant donnée une mesure σ-finie λ sur un espace mesurable $(E, \mathcal{E})$ on dit que N est une mesure de Poisson d'intensité λ si N est une variable aléatoire à valeurs dans $\mathcal{M}_p(E)$ telle que*

- *si les $(A_i)_{i\in I}$ forme une famille d'ensembles mesurables deux à deux disjoints, alors les variables aléatoires $(N(A_i))_{i\in I}$ sont indépendantes.*
- *Si $A \in \mathcal{E}$ vérifie $\lambda(A) < +\infty$, alors la variable aléatoire $N(A)$ suit une loi de Poisson de paramètre $\lambda(A)$: $N(A) \sim \mathcal{P}(\lambda(A)$.*

Un processus de Poisson peut également être considéré comme un nuage aléatoire de points D_ω, et alors $N(A) = \sum_{x\in D_\omega} \mathbf{1}_{(x\in A)}$ est la variable aléatoire qui compte le nombre de points qui tombent dans A.

Théorème 1.7. — *Etant donnée une mesure σ-finie λ sur un espace mesurable $(E, \mathcal{E})$, il existe un processus de Poisson d'intensité λ.*

Démonstration. — Supposons dans un premier temps que λ est une mesure finie : on pose $a = \lambda(E) < +\infty$ et on considère:

- M une variable aléatoire de loi de Poisson de paramètre a : $M \sim \mathcal{P}(a)$.
- $(X_n)_{n\in\mathbb{N}}$ indépendantes de même loi $P_X = \frac{1}{a}\lambda$, et indépendantes de la variable aléatoire M.

On pose enfin :

$$N = N(\omega) = \sum_{i=1}^{M(\omega)} \delta_{X_i(\omega)} .$$

Montrons dans un premier temps que $N(A) \sim \mathcal{P}(\lambda(A))$. On a $N(A) = \sum_{i=1}^{M} \mathbf{1}_{(X_i\in A)}$. C'est un exercice classique : on peut par exemple calculer la fonction génératrice: si $u > 0$, en raison de l'indépendance de M et de la suite $(X_n)_{n\in\mathbb{N}}$, on a

$$\begin{aligned}
\mathbb{E}\left[u^{N(A)}\right] &= \sum_k \mathbb{E}\left[u^{N(A)} \mid M = k\right] \mathbb{P}(M = k) \\
&= \sum_k \mathbb{E}\left[u^{\mathbf{1}_{(X_1\in A)}+\cdots+\mathbf{1}_{(X_k\in A)}}\right] \mathbb{P}(M = k) \\
&= \sum_k \mathbb{E}\left[u^{\mathbf{1}_{(X_1\in A)}}\right]^k \mathbb{P}(M = k) \\
&= \sum_k v^k \mathbb{P}(M = k) = \mathbb{E}\left[v^M\right] = e^{-a(1-v)} = e^{-\lambda(A)(1-u)}
\end{aligned}$$

car $v = \mathbb{E}\left[u^{\mathbf{1}_{(X_1\in A)}}\right] = u\mathbb{P}(X_1 \in A) + 1 - \mathbb{P}(X_1 \in A) = 1 - (1-u)\lambda(A)/a$. Pour démontrer que l'on a réellement une mesure de Poisson, on utilise la caractérisation suivante, appelée également “Master Formula”

Proposition 1.8. — *La mesure ponctuelle aléatoire N est une mesure de Poisson d'intensité λ si et seulement si pour toute fonction $f : E \to [0, +\infty]$ mesurable,*

$$\mathbb{E}\left[e^{-N(f)}\right] = exp\left(-\int_E (1 - e^{-f(x)})\, d\lambda(x)\right).$$

On peut alors utiliser pratiquement le même argument que précédemment:

$$\begin{aligned}
\mathbb{E}\left[e^{-N(f)}\right] &= \sum_k \mathbb{E}\left[e^{-N(f)} \mid M = k\right] \mathbb{P}(M = k) \\
&= \sum_k \mathbb{E}\left[e^{-(f(X_1)+\cdots f(X_k))}\right] \mathbb{P}(M = k) \\
&= \sum_k \mathbb{E}\left[e^{-f(X_1)}\right]^k \mathbb{P}(M = k) \\
&= \exp(-a(1 - \mathbb{E}\left[e^{-f(X_1)}\right]))
\end{aligned}$$

et on conclut car:

$$a(1 - \mathbb{E}\left[e^{-f(X_1)}\right]) = a\mathbb{E}\left[1 - e^{-f(X_1)}\right] = a \int \frac{1}{a}(1 - e^{-f(x)}) d\lambda(x) = \int_E (1 - e^{-f(x)})\, d\lambda(x)$$

Traitons maintenant le cas où la mesure λ n'est pas finie. Comme elle est σ-finie, on peut partitionner $E = \cup_{p \in \mathbb{N}} E_p$ avec les E_p deux à deux disjoints et $\lambda(E_p) < +\infty$. On construit alors N_p mesure de Poisson, à valeurs dans E_p, d'intensité $\lambda_p = \lambda_{E_p}$ restriction de λ à E_p. L'important est de construire ces mesures E_p de façon indépendante les unes des autres. Il suffit ensuite de poser, pour f mesurable positive,

$$N(f) = \sum_p N_p(f_{|E_p})$$

Les règles d'additivité des paramètres pour des variables de Poisson indépendantes entrainent alors automatiquement que $N(f)$ suit une loi de Poisson de paramètre

$$\sum_p \lambda_p(f_{|E_p}) = \lambda(f).$$

□

de la Proposition 1.8. — $\Leftarrow$ Si les $(A_i)_{i\in I}$ forme une famille finie d'éléments 2 à 2 disjoints, alors on pose $f = \sum_i a_i 1_{A_i}$, pour des réels $a_i \geq 0$ et on obtient:

$$\begin{aligned}
\mathbb{E}\left[e^{-N(f)}\right] = \mathbb{E}\left[e^{-\sum_i a_i N(A_i)}\right] &= \exp - \int (1 - e^{-f(x)})\, d\lambda(x) \\
&= \exp - \sum_i \int_{A_i} (1 - e^{-a_i})\, d\lambda(x) \\
&= \exp - \sum_i (1 - e^{-a_i})\lambda(A_i) = \prod_i \mathbb{E}\left[e^{-a_i N(A_i)}\right]
\end{aligned}$$

ce qui prouve en considérant un seul indice que $N(A_i)$ suit bien une loi de Poisson de paramètre $\lambda(A_i)$, puis en considérant tous les indices que les variables aléatoires $(N(A_i))_{i\in I}$ sont indépendantes.
$\Rightarrow$ Pour la réciproque, en reprenant le calcul ci dessus à l'envers, on remarque que l'on a démontré la formule pour f étagée, puis on utilise le fait qu'une fonction mesurable positive f est limite croissante de fonctions simples f_n , et on passe à la limite dans les deux côtés de la formule:

$$N(f) = \lim \uparrow N(f_n), \quad (1 - e^{-f}) = \lim \uparrow (1 - e^{-f_n})\,.$$

$\square$

En utilisant la construction de la preuve précédente ainsi que l'unicité en loi d'une mesure de Poisson d'ntensité donnée, nous obtenons les propriétés suivantes.

Proposition 1.9. — *Soit N une mesure de Poisson sur E d'intensité μ.*

1. *Soit $A \in \mathcal{E}$ tel que $\mu(A) < +\infty$. ALors $N(A) \sim \mathcal{P}(\mu(A))$ et conditionnellement à $N(A) = k$, $N_{|A}$ a même loi que $\sum_{i=1}^k \delta_{X_i}$ avec $(X_i)_{i\geq 1}$ IID de loi $\mu(. \mid A)$.*
2. *Si $A_1, A_2, \ldots, A_k \in \mathcal{E}$ sont disjoints alors les restrictions $M_{|A_i}$ sont des mesures de Poisson indépendantes d'intensités $\mu_{|A_i}$.*
3. *Tout mesure de Poisson peut s'écrire $N = \sum_{i\in I} \delta_{X_i}$ avec I un ensemble aléatoire fini ou dénombrable, et les X_i des variables aléatoires à valeurs dans E (pas nécéssairement IID)*
4. *Si $f \geq 0$ mesurable ou $f \in L^1(\mu)$ alors*

$$\mathbb{E}[N(f)] = \int f\, d\mu\,.$$

Démonstration. — Seule la dernière identité est à prouver. On suppose $f \geq 0$. On peut soit appliquer la formule exponentielle à αf et dériver par rapport à α en prenant $\alpha = 0$, soit écrire directement $N(f) = \sum_p N_p(f)$ et remarquer que par l'indentité de Wald

$$N_p(f) = \mathbb{E}\left[\sum_{i=1}^{M_p} f(X_i^{(p)})\right] = \mathbb{E}\left[f(X_1^{(p)})\right]\mathbb{E}\left[M_p\right] = \mu_p(f)\mu(E_p) = \int_{E_p} f\, d\mu$$

et ensuite sommer ces identités.

□

Il est maintenant facile de démontrer les proriétés de stabilité des mesures de Poisson.

***Lemme 1.10* (Image).** — *Soit m une mesure de Poisson d'itensité μ sur $(E, \mathcal{E})$ et $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ mesurable. Si la mesure image $f * \mu$ est σ-finie, lors l'image $f * M$ est une mesure de Poisson sur F d'intensité $f * \mu$.*

Démonstration. — On a $M = \sum_{1\leq i\leq N} \delta_{X_i}$ et donc $f * M = \sum_{1\leq i\leq N} \delta_{f(X_i)}$. Pour tout ϕ mesurable positive
(1)

$$\mathbb{E}\left[e^{-f*M(\phi)}\right] = \mathbb{E}\left[e^{-M(\phi\circ f)}\right] = \exp(-\int(1-e^{-\phi\circ f(x)})d\mu(x) = \exp\left(-\int(1-e^{-\phi(y)})df*\mu(y)\right)$$

□

Remarque. — *On remarque que si $f * \mu$ n'est pas sigma-finie, cela entraîne que pour toute ϕ mesurable, $\phi > 0$ on a $\int \phi df * \mu = +\infty$, et donc il y a explosion car $f * M(\phi) = +\infty$ p.s. et donc ps $f * M$ n'est pas σ-finie.*

***Lemme 1.11* (Superposition).** — *Soit M_1, M_2 deux mesures de Poisson sur E indépendantes, d'intensités μ_1, μ_2. Alors $M_1 + M_2$ est une mesure de Poisson d'intensité $\mu_1 + \mu_2$.*

La preuve est triviale.

***Lemme 1.12* (Marquage).** — *Soit $M = \sum_{i=1}^N \delta_{X_i}$ une mesure de Poisson d'intensité μ et soit $(Y_n)_{n\in\mathbb{N}}$ une suite IID de variable aléatoire de loi ν sur $(F, \mathcal{F})$, indépendante de M c'est à dire indépendante de $\sigma(M) = \sigma(N, x_i, i \geq 1)$. Alors $M' := \sum_{1\leq i\leq N} \delta_{(X_i,Y_i)}$ est une mesure de Poisson sur $E \times F$ d'intensité $\mu \otimes \nu$.*

Démonstration. — Soit $f : E \times F \to \mathbb{R}$ mesurable positive

$$\begin{aligned}\mathbb{E}\left[e^{-M'(f)} \mid \sigma(M)\right] &= \mathbb{E}\left[\prod_{i=1}^N e^{-f(X_i,Y_i)} \mid \sigma(M)\right] \\ &= \prod_{i=1}^N \int e^{-f(X_n,y)}\, d\nu(y) = e^{-M(F)}\end{aligned}$$

avec $F(x) = log(\int e^{-f(x,y)}\, d\nu(y))$. Donc,

$$\mathbb{E}\left[e^{-M'(f)}\right] = \mathbb{E}\left[e^{-M(F)}\right] = \exp(-\int (1-e^{-F(x)})d\mu(x)) = \exp\left(-\int (1-e^{-f(x,y)})d\mu(x)d\nu(y)\right).$$

□

Lemme 1.13 **(Désintégration/Raréfaction/Thinning)**
Soit $M = \sum_{1\le i\le N} \delta_{X_i}$ *une mesure de Poisson d'inetnsité* μ. *Soit* $(Y_n)_{n\ge 1}$ *une suite IID de variables de loi* $\mathcal{B}(p)$, *indépendante de* M. *Soit*

$$M_1 = \sum_{1\le i\le N, Y_i=1} \delta_{X_i} \quad M_0 = \sum_{1\le i\le N, Y_i=0} \delta_{X_i}$$

Alors M_1 *et* M_0 *sont deux mesures de Poisson indépendantes d'intensités* $p\mu$ *et* $(1-p)\mu$.

Remarque. — *Cela signifie que si l'on reçoit les clients à la porte d'une boutique, qu'il y en a en moyenne* λ *par unité de temps, et qu'on les dispatche avec la proba* p *sur le vendeur 1 et la proba* $1-p$ *sur le vendeur 2, alors le vendeur 1 verra arriver un flux poissonnien de clients, en moyenne* $p\lambda$ *par unité de temps (et que les flux arrivant sur les deux vendeurs sont indépendants).*

Démonstration. — On considère $M' = \sum_{1\le i\le N} \delta_{(X_i,Y_i)} \sim \mathcal{P}(\nu = \mu \times Ber(p))$. On le décompose suivant deux ensembles mesurables disjoints $A_1 = E \times \{1\}$, $A_0 = E \times \{0\}$. Cela donne deux mesures de Poisson indépendantes M'_i d'intensités $\nu_{|A_i}$. Puis on remarque que $M_i = f * M'_i$ avec $f(x,y) = x$. DOnc les M_i sont des mesures de Poisson indépendantes d'intensité $f * \nu_{|A_1} = p\mu$ et $f * \mu_{|A_0} = (1-p)\mu$. □

3.1 Construction du processus de Poisson homogène On prend pour μ la mesure de Lebesgue sur $\mathbb{R}^+$ et on pose $N(t) = M(]0,t])$. Alors par la construction précédente la Proposition 1.6 est immédiate, comme sont immédiates les propriétés d'independance et de stationnarité des accroissements : si $A_i =]t_i, t_{i+1}]$ comme $t_1 < t_1 < ... < t_n$ les A_i sont deux a deux disjoints donc les $N(A_i) = N(t_{i+1}) - N(t_i)$ sont indépendantes de loi de Poisson de paramètres $\lambda(A_i) = t_{i+1} - t_i$.
Pour montrer que cette construction donne bien un processus qui a la même loi que le processus de comptage construit à partir d'exponentielles IID, il faut montrer que les intervalles entre les temps de saut sont des exponentielles IID.

Lemme 1.14. — *Le premier temps de saut du processus* N

$$T = \inf\{t > 0 : N_t = 1\}$$

suit une loi exponentielle.

Démonstration. — Comme $\{T > t\} = \{N_t = 0\}$ avec $N_t \sim \mathcal{P}(\lambda t)$ on a

$$\mathbb{P}(T > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$$

□

***Proposition 1.15* (Propriété de Markov).** — *Soit N un Processus de Poisson de paramètre λ. Pour tout $s \geq 0$, $(\bar{N}_t = N_{t+s} - N_s)_{t\geq 0}$ est un Processus de Poisson de paramètre λ. En outre $\bar{N}$ est indépendant du passé avant s : $(N_u, u \leq s)$.*

Démonstration. — Les accroissements de $\bar{N}$ sont juste des translatés de s des accroissements de N: $\bar{N}_{u+v} - \bar{N}_u = N_{u+v+s} - N_{u+s}$ Ils sont donc indépendants et stationnaires. Tout comme N , le processus $\bar{N}$ et croissant, continu à droite et ne croît que par sauts de 1. C'est don un processus de Poisson de paramètre disons $\mu > 0$. Comme $\bar{N}_t = N_{t+s} - N_s \sim \mathcal{P}(\lambda t)$ on a bien $\mu = \lambda$.
Pour montrer que $\bar{N}$ est indépendant du passé avant s : $(N_u, u \leq s)$, il suffit de montrer que pour tous $0 \leq u_1 \leq \ldots \leq u_n \leq s$ et $0 \leq t_1 \leq \ldots \leq t_p$ les vecteurs $(N_{u_1}, \ldots, N_{u_n})$ et $(\bar{N}_{t_1}, \ldots, \bar{N}_{t_p})$ sont indépendants. On écrit ces deux vecteurs en fonction des accroissements de N, avant et après s, ce qui montre immédiatement l'indépendance.

□

***Proposition 1.16* (Propriété de Markov forte).** — *Soit N un Processus de Poisson de paramètre λ. Soit T un temps d'arrêt fini. Pour tout $s \geq 0$, $(\bar{N}_t = N_{t+T} - N_T)_{t\geq 0}$ est un Processus de Poisson de paramètre λ. En outre $\bar{N}$ est indépendant du passé avant T : $(N_u, u \leq T)$.*

On peut maintenant montrer que les variables $\tau_i = T_{i+1} - T_i$ sont IID avec $T_0 = 0$ et $T_i = \inf\{t > 0 : N_t = i\}$. On sait dejà que $\tau_1 \sim \mathcal{E}(\lambda)$. On remarque que $\tau_2 = \tau_1(\tilde{N})$ avec $\tilde{N}_s = N_{t+\tau_1} - N_{\tau_1} = N_{t+\tau_1} - 1$. D'après la proposition précédente τ_2 est indépendant de $(N_u, u \leq \tau_1)$ donc de τ_1 et a même loi que τ_1. On répète l'opération avec τ_3 indépendnant de τ_1, τ_2 ...

3.2 Le paradoxe de l'autobus On suppose que les intervalles de passage $X_1, \ldots, X_n, \ldots$ des autobus à l'arrêt Morrhonière sont des variables indépendantes de même loi $exp(\lambda)$. Le nombre de bus passé avant t est donc $N_t = \sum_{n\geq 1} \mathbf{1}_{(S_n \leq t)}$ avec $S_n = X_1 + \cdots + X_n$. C'est un Processus de Poisson de paramètre λ. Sachant que l'on arrive à l'arrêt à l'instant $t > 0$:

- Le temps d'attente du prochain bus est

$$R_t = \inf\{s > t : N_s = N_t + 1\} = S_{1+N_t} - t$$

- Le temps écoulé depouis le passage du dernier bus est

$$A_t = \sup\{s \leq t : N_s < N_t\} = t - S_{N_t}$$

Proposition 1.17. — *Les variables aléatoires R_t et A_t sont indépendantes, $R_t \sim exp(\lambda)$ et $A_t \overset{d}{=} \inf(\mathbf{e}, t)$ avec $\mathbf{e} \sim exp(\lambda)$.*

Démonstration. — A_t ne dépend que de $(N_s, s \leq t)$ et R_t ne dépend que de $(\bar{N}_s = N_{t+s} - N_t, s \geq 0)$. Par la propriété de Markov, A_t et R_t sont indépendantes.

$$R_t = \inf \left\{ s \geq 0 : \bar{N}_s = 1 \right\}$$

est le premier temps de saut de $\bar{N}$, donc $R_t \sim exp(\lambda)$. On a $0 \leq A_t \leq t$ et si $0 < s < t$, alors

$$\{A_t \geq s\} = \left\{ N'_s = 0 \right\}$$

avec N' le Processus de Poisson de paramètre λ défini par $N'_u = N_{t-s+u} - N_{t-s}$. Donc,

$$\mathbb{P}(A_t \geq s) = \mathbb{P}\left(N'_s = 0\right) = e^{-\lambda s}.$$

D'autre part $\mathbb{P}(A_t = t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$ ce qui prouve que $A_t \overset{d}{=} \inf(\mathbf{e}, t)$.

□

Remarque. — *On a $\mathbb{E}[R_t] = \frac{1}{\lambda}$ et $\mathbb{E}[A_t] \to \mathbb{E}[\mathbf{e}] = \frac{1}{\lambda}$ par convergence monotone. Donc si $L_t = R_t + A_t$ est la longueur de l'intervalle dans lequel on tombe on obtient $\mathbb{E}[L_t] \to 2/\lambda$. C'est un paradoxe : si les autobus passent en moyenne toutes les 10 minutes, ie $\frac{1}{\lambda} = 10$, alors l'intervalle dans lequel on tombe a une longueur moyenne très proche de 20 minutes. En outre, on attendra en moyenne $\mathbb{E}[R_t] = \frac{1}{\lambda}$=10 minutes le prochain bus, et ce indépendamment du temps écoulé depuis le passage du bus précédent.*

SECTION 4

Le processus de Poisson composé

Exemple 1.2. — *Le nombre de voitures qui s'arrêtent entre midi et treize heures devant le Mcdo de Rezé suit une loi de Poisson de paramètre $\lambda = 180$. Le nombre de passagers par voiture suit une loi binômiale $\mathcal{B}(n = 5, p = \frac{1}{2})$. Calculer la moyenne et la variance du nombre de clients X.*
Le bon sens dit que $\mathbb{E}[X] = 180 \times 2,5 = 450$. Mais comment calculer la variance ? Il est naturel de supposer que si Y_i est le nombre de passagers de la i-ème voiture, alors les variables aléatoires Y_i sont indépendantes de même loi $\mathcal{B}(n = 5, p = \frac{1}{2})$. Si N_t est le nombre de voitures arrivées avant $t = 1$ heure, on commence l'échelle des temps à midi, on a

$$X = \sum_{i=1}^{N_t} Y_i .$$

Définition 1.5. — *Un processus stochastique $(X_t)_{t\geq 0}$ est appelé Processus de Poisson composé s'il peut être représenté par $X_t = \sum_{i=1}^{N_t} Y_i$ avec $(N_t)_{t\geq 0}$ un Processus de Poisson indépendant de la suite IID $(Y_i)_{i\geq 1}$.*

La proposition suivante permet de calculer aisément espérance et variance d'un processus de Poisson composé de paramètre λ.

Proposition 1.18. — *1. Si Y_1 est intégrable alors X_t est intégrable et on a $\mathbb{E}[X_t] = \lambda t \mathbb{E}[Y_1]$.*
2. SI Y_1 est de carré intégrable, alors X_t l'est aussi et $\mathrm{Var}(X_t) = \lambda t \mathbb{E}\left[Y_1^2\right]$.

Pour l'exemple du Mcdo, cela donne $\mathrm{Var}(X) = 180 \times 7,5 = 1350$. Ce qui donne une très forte probabiité d'observer entre 300 et 600 personnes. La Proposition est une conséquence immédiate du

Lemme 1.19. — *Soit $X = \sum_{i=1}^{N} Y_i$ avec N une variable aléatoire entière indépendante de la suite IID $(Y_i)_{i\geq 1}$.*
1. *Si N et Y sont intégrables, alors on a l'identité de Wald: $\mathbb{E}[X] = \mathbb{E}[N]\mathbb{E}[Y]$.*
2. *Si N et Y sont dans $\mathcal{L}^2$ alors $X \in \mathcal{L}^2$ et*
$$\mathrm{Var}(X) = \mathbb{E}[N]\,\mathrm{Var}(Y) + \mathrm{Var}(N)\mathbb{E}[Y]^2 .$$

Démonstration. — 1. Si Y_1 était une variable aléatoire entière on pourrait raisonner avec les fonctions génératrices. Cependant, le cas général n'est pas plus compliqué:

$$\begin{aligned}
\mathbb{E}[|X|] &= \sum_n \mathbb{P}(N=n)\mathbb{E}[|X| \mid N=n] \\
&= \sum_n \mathbb{P}(N=n)\mathbb{E}[|Y_1 + \cdots Y_n|] \\
&\leq \sum_n \mathbb{P}(N=n)(\mathbb{E}[|Y_1|] + +\mathbb{E}[|Y_1|]) \\
&= \mathbb{E}[|Y_1|]\sum_n \mathbb{P}(N=n)n = \mathbb{E}[|Y_1|]\mathbb{E}[N] < +\infty
\end{aligned}$$

donc X est intégrable et un calcul analogue, sans les valeurs absolues, montre que $\mathbb{E}[X] = \mathbb{E}[Y]\mathbb{E}[N]$.

2. On calcule la fonction caractéristique de X

$$\begin{aligned}\phi_X(u) = \mathbb{E}\left[e^{iuX}\right] &= \sum_n \mathbb{P}(N=n)\mathbb{E}\left[e^{iuX} \mid N=n\right] \\ &= \sum_n \mathbb{P}(N=n)\mathbb{E}\left[e^{iu(Y_1+\ldots+Y_n)}\right] \\ &= \sum_n \mathbb{P}(N=n)\phi_{Y_1}(u)^n = \mathbb{E}\left[\phi_{Y_1}(u)^N\right] \\ &= G(\phi_{Y_1}(u)).\end{aligned}$$

Comme Y_1 est de carré intégrable

$$\phi_{Y_1}(u) = 1 + u\mathbb{E}[Y_1] - (u^2/2)\mathbb{E}\left[Y_1^2\right] + o(u^2).$$

Comme N est de carré intégrable,$G'(s)$ et $G''(s)$ admette des limites finies quand $s \uparrow 1$,

$$G(1-a) = 1-aG'(1)+(a^2/2)G''(1)+o(a^2) = 1-a\mathbb{E}[N]+(a^2/2)G''(1)\mathbb{E}[N(N-1)]+o(a^2).$$

et on en déduit un développement limité en 0 d'ordre 2 de $\phi_X(u)$ qui nous donne $\mathbb{E}\left[X^2\right]$.
On peut aussi utiliser la même technique qu'au 1 pour calculer $\mathbb{E}\left[X^2\right]$ directement :

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \sum_n \mathbb{P}(N=n)\mathbb{E}\left[(Y_1+\cdots+Y_n)^2\right] \\ &= \sum_n \mathbb{P}(N=n)(var(Y_1+\cdots+Y_n) + \mathbb{E}[(Y_1+\cdots+Y_n)]^2) \\ &= \sum_n \mathbb{P}(N=n)(n\operatorname{Var}(Y_1) + n^2\mathbb{E}[Y_1]^2) = \mathbb{E}[N]\operatorname{Var}(Y_1) + \mathbb{E}\left[N^2\right]\mathbb{E}[Y_1]^2\end{aligned}$$

$\square$

Remarque. — *On peut observer que si $M = \sum_{i=1}^N \delta_{\xi_i}$ est une mesure de Poisson d'intensité λ et $N_t = M(]0,t])$ alors $X_t = M'(f_t)$ avec M' le processus marqué*

$$M' = \sum_{1\leq i\leq N} \delta_{(\xi_i, Y_i)}$$

et $f_t(\xi, y) = \mathbf{1}_{(\xi\leq t)}y$. Alors il vient, si ν est la loi de Y_1

$$\mathbb{E}[X_t] = \int f_t(\xi,y)\lambda d\xi d\nu(y) = \lambda t \int y d\nu(y) = \lambda t\mathbb{E}[Y]$$

De même on calcule la transformée de Laplace

$$\mathbb{E}\left[e^{-\alpha X_t}\right] = \mathbb{E}\left[e^{-M(\alpha f_t)}\right] = \exp\left(-\int(1-e^{-\alpha f_t(\xi,y)})\lambda d\xi d\nu(y)\right) = \exp\left(-\lambda t\mathbb{E}\left[1-e^{-\alpha Y}\right]\right)$$

et on en déduit avec $\psi(\alpha) = -\log \mathbb{E}\left[e^{-\alpha X_t}\right]$*,* $\mathbb{E}\left[X_t\right] = \psi'(0)$ *et* $\mathrm{Var}(X_t) = \psi''(0) = \lambda t \mathbb{E}\left[Y^2\right]$.

SECTION 5

Le processus de Poisson inhomogène

C'est un Processus de Poisson d'intensité variable. Soit $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction borélienne positive localement intégrable, i.e. pour tous $0 < a < b < +\infty$ $\int_a^b \lambda(t)\, dt < +\infty$.

Définition 1.6. — *On dit que* $(N_t)_{t\geq 0}$ *est un Processus de Poisson inhomogène d'intensité* λ *si* $(N_t)_{t\geq 0}$*est un processus à accroissements indépendants et pour tous* $s, t \geq 0$

$$N_{t+s} - N_s \sim \mathcal{P}\left(\int_s^{t+s} \lambda(u)\, du\right).$$

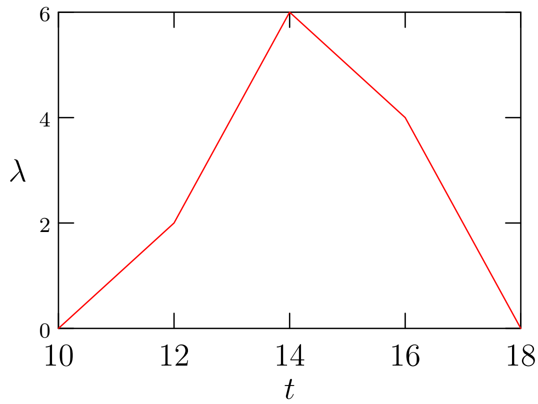


Figure 1. Intensité du processus d'arrivée des clients

Exemple 1.3. — *On observe le taux d'arrivée, exprimé en clients par heure, à la boutique de harry. Il est* 0 *à 10 heures,* 4 *à midi,* 6 *à 14 heures,* 2 *à 16 heures et* 0 *à 18 heures. Entre ces valeurs il varie linéairement. Faire un dessin. On désire déterminer la loi du nombre de clients dans une journée et la probabilité que personne n'arrive avant midi.*
Pour cela on suppose donc que le nombre de clients suit un Processus de Poisson inhomogène de taux λ *linéaire par morceaux. Le nombre de clients arrivés dans la journée est donc*

$$C = N_{18} - N_{10} \sim \mathcal{P}\left(\int_{10}^{18} \lambda(t)\, dt\right) = \mathcal{P}(20).$$

On calcule l'intégrale en calculant l'aire sous la courbe, i.e. en additionnant l'aire des triangles ! De même le nombre de clients arrivés avant midi est $M = N_{12} - N_{10} \sim \mathcal{P}(4)$ et donc $\mathbb{P}(M = 0) = e^{-4}$.
En exercice, on peut, en supposant que Harry ferme exceptionnellement sa boutique à 17h30, calculer le nombre moyen de clients perdus et la probabilité qu'au moins un client trouve porte close (on trouve 0,5 et 0,39).

Observons que l'on ne peut pas savoir si on a affaire à un Processus de Poisson inhomogène. On fait une hypothèse de modélisation qui ne peut être validée que par une étude statistique.
Il et très facile de construire un Processus de Poisson inhomogène. Il suffit de changer l'horloge.

Proposition 1.20. — *Si $(N_t)_{t\geq 0}$est un Processus de Poisson d'intensité 1, et $F(t) = \int_0^t \lambda(s)\, ds$, alors $M_t = N_{F(t)}$ est un Processus de Poisson inhomogène d'intensité λ.*

Une fois faite l'hypothèse de modélisation par un Processus de Poisson inhomogène, on identifie la fonction intensité à l'aide du

Lemme 1.21. — *Soit $(N_t)_{t\geq 0}$un Processus de Poisson inhomogène d'intensité λ, alors si λ est continue à droite en t:*

$$\mathbb{P}(N_{t+h} - N_t \neq 0) = h\lambda(t) + o(h)\,.$$

Démonstration. — EN effet, t étant fixé, la variable aléatoire $X = N_{t+h} - N_t$ suit une loi de Poisson de paramètre $f(h) = \int_t^{t+h} \lambda(s)\, ds = h\lambda(t) + o(h)$ et $\mathbb{P}(X \neq 0) = 1 - e^{-f(h)} = h\lambda(t) + o(h)\,.$ □

SECTION 6

Universalité de la loi de Poisson

Nous allons généraliser le théorème de convergence de la loi binômiale vers la loi de Poisson :

$$\text{si } np_n \to \lambda > 0\,, \quad \text{alors} \quad \mathcal{B}(n, p_n) \Rightarrow \mathcal{P}(\lambda).$$

Théorème 1.22. — *Soit $(X_{n,m})_{1\leq m\leq n}$ des variables aléatoires entières indépendantes telles que si $p_{n,m} = \mathbb{P}(X_{n,m} = 1)$ et $\delta_{n,m} = \mathbb{P}(X_{n,m} \geq 2)$ alors*

(i) $\sum_{1\leq m\leq n} p_{n,m} \underset{n\to\infty}{\longmapsto} \lambda > 0$

(ii) $\sup_{1\leq m\leq n} p_{n,m} \underset{n\to\infty}{\longmapsto} 0$

(iii) $\sum_{1\leq m\leq n} \delta_{n,m} \to 0$

Alors la suite $S_n = X_{n,1} + \cdots + X_{n,}$, converge en loi vers une variable aléatoire Z qui suit une loi de Poisson de paramètre $\lambda > 0$.

Application : un système de marcheurs aléatoires Dans une rue, supposée très grande, on place des clients, qui se promènent au hasard. Si on compte le nombre de clients qui sont à un instant donné, assez grand, devant une boutique précise, on observe une variable aléatoire qui suit une loi de Poisson.
Le modèle mathématique est le suivant. La rue est $\mathbb{Z}$. A l'origine des temps $n = 0$, on place un client en chaque site x de $\mathbb{Z}$: S_n^x est la position à l'instant n du marcheur qui à l'origine des temps était en x, $S_x^0 = x$. La boutique est au point 0 de Z. Le nombre de clients devant la boutique à l'instant n est donc

$$Y_n = \sum_{x \in \mathbb{Z}} \xi_{n,x} \quad \text{avec} \quad \xi_{n,x} = \mathbf{1}_{(S_n^x = 0)}.$$

On désire montrer que $Y_n \to Z \sim \mathcal{P}(1)$. En fait la somme est restreinte à une somme finie de variables aléatoires indépendantes qui prennent les valeurs 0 et 1 : $Y_n = \sum_{|x| \leq n} \xi_{n,x}$. Pour appliquer le théorème 1.22, il reste à montrer que

$$\sum_{|x| \leq n} \mathbb{P}(\xi_{n,x} = 1) \to 1 \quad \text{et} \quad \sup_{|x| \leq n} \mathbb{P}(\xi_{n,x} = 1) \to 0\,.$$

Qu'est ce donc qu'un marcheur issu de x ? Sa position à l'instant n est

$$S_n^x = x + \eta_1 + \cdots + \eta_n$$

avec les variables aléatoires η_i indépendantes de même loi $\mathbb{P}(\eta_i = \pm 1) = \frac{1}{2}$. En observant que $\frac{1}{2}(1 + \eta_i) \sim \mathcal{B}(1, \frac{1}{2})$ on voit que $S_n^x = x + 2B_n - n$ avec $B_n \sim \mathcal{B}(n, \frac{1}{2})$ et donc que

$$\mathbb{P}(\xi_{n,x} = 1) = \mathbb{P}(x + 2B_n - n = 0) = \mathbb{P}\left(B_n = \frac{n - x}{2}\right).$$

Donc

$$\sum_{|x| \leq n} \mathbb{P}(\xi_{n,x} = 1) = \sum_{|x| \leq n} \mathbb{P}\left(B_n = \frac{n - x}{2}\right) = \sum_k \mathbb{P}(B_n = k) = 1$$

et évidemment

$$\sup_{|x| \leq n} \mathbb{P}(\xi_{n,x} = 1) = \sup_k \mathbb{P}(B_n = k) \to 0\,.$$

Remarque. — *On peut même démontrer que si on a une boutique à chaque site $x \in \mathbb{Z}$,alors pour n grand, devant chaque boutique il y a un nombre de clients suivant une loi $\mathcal{P}(1)$ et que ces variables aléatoires sont toutes indépendantes et de moyenne 1.*

2

Renouvellement

SECTION 1

Processus de renouvellement

On considère un processus de comptage $(N(t), t \geq 0)$ dans lequele les intervalles de temps entre les arrivées sont encore IID mais pas forcément exponentiels. Soit $(X_n, n \geq 1)$ IID positives de moyenne $\mu \in (0, +\infty)$, $S_0 = 0$ et $S_n = X_1 + \cdots + X_n$ l'époque de la nième arrivée.

Définition 2.1. — *$N(t) = \sum_{n \geq 1} \mathbf{1}_{(S_n \leq t)} = \sup\{n \geq 1 : S_n \leq t\}$ avec $\sup \emptyset = 0$) est le processus de renouvellement engendré par la suite $(X_i, i \geq 1)$.*

Exemple 2.1. — *Inventaire. La demande du produit pendant la semaine $t = 1, 2, \ldots$ est une suite IID $(X_t, t \geq 1)$. Alors $1 + N(u)$ est le nombre de semaines avant l'épuisement du stock $u \geq 0$.*

***Théorème 2.1* (loi des grands nombres).** —

$$\lim_{t \to +\infty} \frac{N(t)}{t} = \frac{1}{\mu} \quad p.s.$$

Démonstration. — Il suffit de remarquer que $\frac{S_n}{n} \to \mu$ et d'utiliser la monotonie de $n \to S_n$ pour avoir l'encadrement

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{1+N(t)}}{N(t)}$$

qui donne le résultat car $N(t) \to +\infty$. □

Proposition 2.2. — *La fonction de renouvellement* $M(t) = \mathbb{E}[N(t)] = \sum_{n\geq 1} \mathbb{P}(S_n \leq t)$ *vérifie*

$$\lim_{t\to+\infty} \frac{M(t)}{t} = \frac{1}{\mu}$$

***Lemme 2.3* (Identité de Wald).** — *Soit* τ *un temps d'arrêt, i.e. pour tout* n, $\{\tau \leq n\}$ *est* $\mathcal{F}_n$ *mesurable. Si* τ *est intégrable, alors*

$$\mathbb{E}[S_\tau] = \mu\mathbb{E}[\tau].$$

Démonstration. — Par conditionnement,

$$\mathbb{E}[S_\tau] = \sum_n \mathbb{E}\left[S_n \mathbf{1}_{(\tau=n)}\right] = \mathbb{E}\left[\sum_{m\geq 1} X_m \mathbf{1}_{(\tau\geq m)}\right]$$

Or $\{\tau \geq m\}^C = \{\tau < m\} \in \mathcal{F}^X_{m-1}$ est indépendant de X_m donc

$$\mathbb{E}[S_\tau] = \sum_{m\geq 1} \mathbb{E}[X_m]\mathbb{E}\left[\mathbf{1}_{(tau\geq m)}\right] = \mu\mathbb{E}[\tau].$$

Remarquons que nou s n'utilisons pas l'intégrabilité de tau car les X_i sont ositives. Si les X_i sont signées, alors il faut utiliser l'intégrabilité. □

Lemme 2.4. — *Si* X_1 *est majorée par* A, *alors,* $\frac{t}{\mu} \leq \mathbb{E}[N(t)] + 1 \leq \frac{t+A}{\mu}$.

Démonstration. — On applique l'identité de Wald à $\tau = \inf\{n : S_n > t\} = 1 + N(t)$ et on obtient

$$\mathbb{E}[S_\tau] = \mu(1 + \mathbb{E}[N(t)]) \geq t.$$

Comme les X_i sont majorées par A on a $1 + N(t) \leq t + A$ d'où la seconde inégalité. □

Démonstration. — (Preuve de la proposition 2.2) Si X_1 est bonréne par A on passe à la limite dans l'encadrement. Dans le cas général on applique le résultat à $\bar{X}_i = X_i \wedge A$ et on a $\bar{S}_n \leq S_n$ donc $\bar{N}(t) \geq N(t)$ et on obtient en faisant $A \to +\infty$)

$$\limsup \frac{M(t)}{t} \leq \lim \frac{\bar{M}(t)}{t} = \frac{1}{\mathbb{E}[X_1 \wedge A]} \to \frac{1}{\mu}$$

□

Exemple 2.2. — *Bill a une console. La durée d'un pack de piles est uniformément répartie entre 30 et 60 jours. Combien de pack de piles utilisera-t-il en 10 ans ? (Une année veut 360 jours).*
On a $\mu = \mathbb{E}[X_1] = 45$ *jours car* $X_1 \sim \mathcal{U}([30, 60])$. *On obtient pour* $t = 3600$ *jours a peu près*

$$80 = t/\mu \leq 1 + M(t) \leq (t + A)/\mu = 81.37$$

donc il utilise entre 80 *et* 81 *packs en moyenne (et* 80 *asymptotiquement)*

Le temps qu'il reste à vivre à l'instant t, temps d'attente de la prochaine arrivée d'unévènement important, est $R_t = S_{1+N(t)} - t$. On prouve, à l'aide de l'équation de renouvellement, le résultat suivant sur l'asymptotique de R_t

Proposition 2.5. — *La suite R_t converge en loi vers R_∞ de densité $\frac{\mathbb{P}(X_1>t)}{\mathbb{E}[X_1]}$. En particulier on a*

$$\mathbb{E}[R_\infty] = \frac{\mathbb{E}\left[X_1^2\right]}{2\mathbb{E}[X_1]}, \quad \mathbb{E}\left[R_\infty^2\right] = \frac{\mathbb{E}\left[X_1^3\right]}{3\mathbb{E}[X_1]}. \tag{2}$$

Pour une loi exponentielle $X_1 \sim \mathcal{E}(\lambda)$ on a $R_\infty \overset{d}{=} X_1$ car c'est un processus de Poisson. Pour une loi uniforme sur $[2a, 3a]$ on a $\mathbb{E}[X] = 2a$ et $\operatorname{Var}(X) = \operatorname{Var}(a(1+2U)) = 4a^2 \operatorname{Var}(U) = \frac{a^2}{3}$ et donc $\mathbb{E}[R_\infty] = \frac{13}{12}a$ est bien plus petite que $\mathbb{E}[X] = 2a$.

SECTION 2

Renewal reward processes

C'est le terme anglais pour unprocessus de renouvellement composé.

Définition 2.2. — *Soit $(N(t), t \geq 0)$ un processus de renouvellement et soit $(R_i, i \geq 1)$ une suite IID de vas réelles intégrables. ALors*

$$R(t) := \sum_{i=1}^{N_t} R_i$$

est un renewal reward process.

Exemple 2.3. — *Une machine est successivement en marche et en panne. $(U_n, n \geq 1)$ est la suite des durées des périodes de marche et $(D_n, n \geq 1)$ est la suite des durées périodes de panne. On note $D(t)$ le temps total passé en panne avant t.*
Le renewal reward process est $R(t)$ associé au processus de renouvellement engendré par $X_i = U_i + D_i$ et de processus de reward $R_i = D_i$. On suppose que $\mathbb{E}[U_1] = 5\mathbb{E}[D_1]$, et on va voir que cela entraîne que $R(t)/t \to \frac{1}{6}$ p.s.
Il doit être clair que quelle que soit la façon de mesure $D(t)$ on a alors $D(t)/t \to \frac{1}{6}$ également car on a

$$R(t) e D(t) \leq R(t) + D_{1+N(t)}$$

***Théorème 2.6* (renewal reward theorem).** — *Si $\mathbb{E}[R_1] < +\infty$ alors*

$$\lim_{t\to+\infty} \frac{R(t)}{t} = \frac{\mathbb{E}[R_1]}{\mathbb{E}[X_1]} \quad p.s.$$

Ce théorème reste valable quelle que soit la façon dont on récupère les récompenses (en début, en fin de cycle ou continûment). On l'écrit *la récomppense moyenne en temps long égale la récompense moyene sur un cycle de renouvellement divisée par la durée d'un cycle de renouvellement*

Démonstration. — Par la loi forte des grands nombres $\frac{1}{n}(R_1 + \cdots + R_n) \to \mathbb{E}[R_1]$ et donc comme $N(t) \to +\infty$ ps, on a

$$\lim_{t\to+\infty} \frac{R(t)}{N(t)} = \mathbb{E}[R_1]$$

Comme $N(t)/t \to \frac{1}{\mu}$ p.s., on a par encadrement,

$$\frac{N(t)}{t}\frac{R(t)}{N(t)} = \frac{R(t)}{t} \to \frac{\mathbb{E}[R_1]}{\mu}.$$

□

***Exemple 2.4* (Coût réel d'une voiture).** — *Supposons que la durée de vie d'une voiture est une variable aléatoire positive $V \sim \mathcal{E}(\lambda)$. Mr Brown achète une voiture neuve dès que la précédente est cassée ou atteint T années. Une voiture neuve coûte A euros, et un coût additionnel de B euros est payé si la voiture tombe définitivement en panne avant T (par exemple pour payer le dépannage, etc ...).*

1. *Quel est le coût moyen par unité de temps de la stratégie de Mr Brown ? Quelle valeur de T doit-il choisir ? A.N. $A = 20000$ euros, $B = 200$ euros, $\lambda = 1/4$.*
2. *On propose un modèle plus réaliste dans lequel entretenir sa voiture pendant n années après l'achat coûte $\alpha n + \beta n^2$ avec $\alpha > 0, \beta > 0$ donnés. Quelle est la valeur optimale de T ? A.N. $\alpha = 300$, $\beta = 1000$.*

On va donner un résultat asymptotique (en temps long). Ici la durée d'un renouvellement/cycle est

$$\mathbb{E}[X_1] = \mathbb{E}[V \wedge T] = \int_0^\infty \mathbb{P}(V \wedge T > t)\, dt = \int_0^T e^{-\lambda t} dt = \frac{1-e^{-\lambda T}}{\lambda} \tag{3}$$

et le coût du cycle est $R_1 = 1 + B\,\mathbf{1}_{(T<V_1)}$ donc

$$\mathbb{E}[R_1] = A + B\mathbb{P}(V > T) = A + Be^{-\lambda T} \tag{4}$$

On obtient finalement comme coût annuel sur le long terme

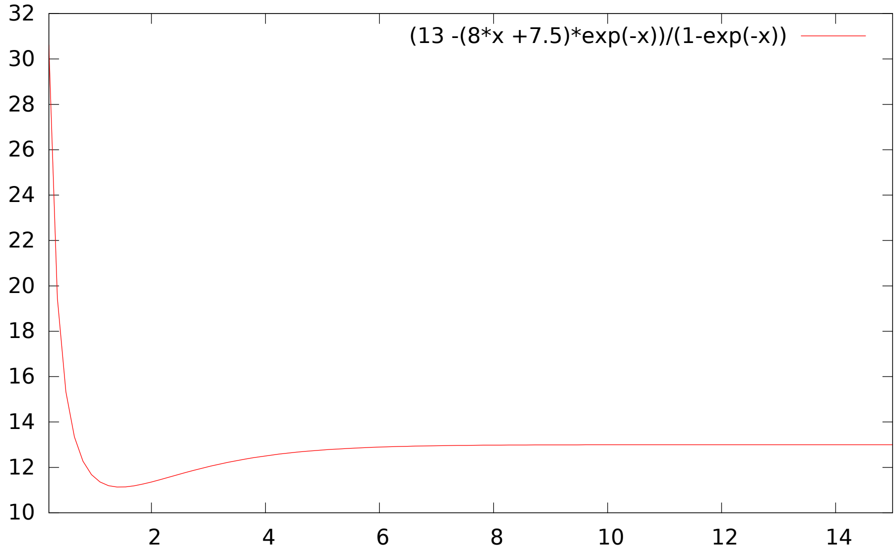
$$c(T) = \lim_{t\to+\infty} \frac{R(t)}{t} = \frac{\mathbb{E}[R_1]}{\mathbb{E}[X_1]} = \lambda\frac{A+Be^{-\lambda T}}{1-e^{-\lambda T}} = \lambda(-B + \frac{A+B}{1-e^{-\lambda T}}).$$

En conséquence la valeur optimale est $T^* = +\infty$ quelles que soient les valeurs de A, B : il ne faut changer la voiture qu'à la première panne, et pas avant.

Pour la seconde question, on voit que l'on obtient comme coût

$$d(T) = c(T) + \alpha + \beta \frac{\mathbb{E}\left[X_1^2\right]}{\mathbb{E}\left[X_1\right]} = -\lambda B + \alpha + \frac{\lambda(A+B) + \frac{2}{\lambda}(1-(1+\lambda T)e^{-\lambda T})}{1-e^{-\lambda T}}$$

Avec les valeurs numériques on voit qu'il faut minimiser la fonction suivante de $x = \lambda T$. On voit que $x^* \approx 1$ et donc $T \approx \frac{1}{\lambda} = 4 = \mathbb{E}\left[V\right]$. Dans ce cas il vaut mieux donc attendre le temps moyen de vie de la voiture!



3

Espérance conditionnelle et loi conditionnelle

SECTION 1

Le cas discret

Motivation 1) On lance un dé à 6 faces. La loi du résultat X est uniforme sur $\{1, \ldots, 6\}$. En fait, on ne voit pas le résultat du dé lancé, mais on nous informe que le résultat est pair. On s'attend à ce que la loi de X devienne une uniforme sur $\{2, 4, 0\}$. Comment formaliser cette intuition ?
2) On observe un grand nombre de patients et on n ote pour chacun d'eux s'il a un taux de cholésterol supérieur à 2 grammes, évènement A, et s'il a pris des statines, évènement B. On désire obtenir une bonne approximation de la probabilité que le patient ait un taux de cholestérol élévé sachant qu'il prend des statines. Il est naturel d'approcher cette probabilité par la proportion $\frac{N_n(A \cap B)}{N_n(B)}$ où A_i (resp B_i) est l'évènement A (resp B pour le i-ème patient), et

$$N_n(A) = \sum_{i=1}^{n} 1_{A_i} \,. \tag{5}$$

Sous l'hypothèse naturelle d'indépendance des patients, on obtient par la loi forte des grands nombres, que presque sûrement,

$$\frac{N_n(A \cap B)}{N_n(B)} = \frac{\frac{1}{n} N_n(A \cap B)}{\frac{1}{n} N_n(B)} \to \frac{\mathbb{P}\,(A \cap B)}{\mathbb{P}\,(B)} \,. \tag{6}$$

Définition 3.1. — *Étant donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et un évènement B de probabilité non nulle la mesure de densité $\mathbb{Q}(\omega) = \frac{1}{\mathbb{P}(B)} 1_B(\omega)\, d\mathbb{P}(\omega)$*

est appelée probabilité conditionnelle sachant B *et on note*

$$\mathbb{P}(A \mid B) = \mathbb{Q}(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \qquad (A \in \mathcal{F}). \tag{7}$$

Il est élémentaire de vérifier alors que si X est une variable aléatoire définie sur $(\Omega, \mathcal{F})$, positive (resp. $\mathbb{P}$-intégrable), alors X est positive (resp. $\mathbb{P}(. \mid B)$) intégrable et

$$\mathbb{E}[X \mid B] = \frac{\mathbb{E}[X 1_B]}{\mathbb{P}(B)}. \tag{8}$$

Définition 3.2. — *Un* système complet d'évènements *est une famille finie ou dénombrable $(B_i)_{i \in I}$ d'évènements deux à deux disjoints, de probabilité non nulle, tels que $\mathbb{P}(\cap_{i \in I} B_i) = 1$ (il y a presque sûrement un et un seul des évènements B_i réalisé).*
On dit que le système d'évènements est quasi complet *si*

$$1 = \sum_{i \in I} 1_{B_i}(\omega) \qquad p.s. \tag{9}$$

Proposition 3.1 **(Formule des probabilités totales)**
Soit X une variable aléatoire positive ou intégrable et $(B_i)_{i \in I}$ un système (quasi) complet d'évènements. Alors

$$\mathbb{E}[X] = \sum_{i \in I} \mathbb{E}[X \mid B_i] \mathbb{P}(B_i), \tag{10}$$

En particulier, pour tout évènement A

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i) \mathbb{P}(B_i). \tag{11}$$

Démonstration. — Il suffit d'utiliser, dans le cas $X \geq 0$, l'interverstion série intégrale pour des fonctions mesurables positives:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[X \sum_i 1_{B_i}\right] = \mathbb{E}\left[\sum_i X 1_{B_i}\right] \\ &= \sum_i \mathbb{E}[X 1_{B_i}] = \sum_{i \in I} \mathbb{E}[X \mid B_i] \mathbb{P}(B_i). \end{aligned}$$

Le cas X intégrable s'en déduit en écrivant $X = X + -X^-$. □

Étant donné un système complet d'évènements, il est naturel de définir *la probabilité conditionnelle de A sachant $(B_i)_{i \in I}$* comme étant la *variable aléatoire* qui vaut $\mathbb{P}(A \mid B_i)$ sur l'évènement B_i. Plus précisément si $B = \sigma(B_i, i \in I)$ est la tribu engendrée par un système complet, alors pour tout évènement A:

$$\mathbb{P}(A \mid \mathcal{B})(\omega) = \sum_{i \in I} \mathbb{P}(A \mid B_i) 1_{B_i}(\omega)\, p.s. \tag{12}$$

Remarque. — *C'est un exercice classique de théorie de la mesure que de vérifier que cette définition de la variable aléatoire $\mathbb{P}(A \mid \mathcal{B})$, à un ensemble de mesure nulle près, ne dépend que de la tribu $\mathcal{B}$ et pas du système complet d'évènements qui l'engendre.*
En effet si $(B_i)_{i\in I}$ est une partition alors $\mathcal{B} = \mathcal{P}$ avec

$$\mathcal{P} = \{A : \exists J \subset I, A = \cap_{i\in J} B_i\}\,. \tag{13}$$

Dans le cas général, si $N = (\cap_{i\in I} B_i)^C$ alors $\mathbb{P}(N) = 0$ et

$$\mathcal{B} = \mathcal{P} \cap \{A \cap N, A \in \mathcal{P}\}\,. \tag{14}$$

Si X est une variable aléatoire positive ou intégrable, on définit de façon analogue

$$\mathbb{E}[X \mid \mathcal{B}](\omega) = \sum_{i \in I} \mathbb{E}[X \mid B_i] 1_{B_i}(\omega) \tag{15}$$

Le cas le plus classique est celui où l'on dispose d'une variable aléatoire Y à valeurs dans un espace discret E ; alors, pour X positive ou intégrable,

$$\mathbb{E}[X \mid Y] = \sum_{y} \mathbb{E}[X \mid Y = y]\, \mathbf{1}_{(Y=y)}\,. \tag{16}$$

Hélas tout se complique lorsque la variable aléatoire Y prend ses valeurs dans un ensemble continu (non discret), car alors $\{Y = y\}$ est un évènement de probabilité nulle. Il est naturel de penser qu'il suffit alors de définir $\mathbb{E}[X \mid Y] = \phi(Y)$ avec $\phi(y) = \mathbb{E}[X \mid Y = y] = \lim \mathbb{E}[X \mid A_n]$ avec A_n une suite d'évènement de probabilité non nulle convergeant vers l'évènement $\{Y = y\}$. Hélas, mille fois hélas, le résultat dépend fortement du choix de la suite A_n, et il n'y a pas de choix canonique d'une telle suite. Un exercice classique sur les couples de gaussiennes dans le plan, que vous trouverez en fin de chapitre, devrait achever de vous convaincre.
Nous allons donc utiliser une généralisation de la caractérisation suivante de l'espérance conditionnelle. La variable X étant intégrable et la variable aléatoire Y discrète, $\mathbb{E}[X \mid Y]$ est l'unique, presque sûrement, variable aléatoire $\phi(Y)$ intégrable telle que pour toute fonction mesurable bornée f:

$$\mathbb{E}[Xf(Y)] = \mathbb{E}[\phi(Y)f(Y)]\,. \tag{17}$$

Exemple 3.1. — *Soit $X \sim \mathcal{E}(\lambda)$ une variable aléatoire exponentielle. Alors X est sans mémoire:*

(18)

$$\mathbb{P}(X \geq t+s \mid X \geq s) = \frac{\mathbb{P}(X \geq t+s)}{\mathbb{P}(X \geq s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X \geq t). \qquad (t, s \geq 0).$$

SECTION 2

Espérance conditionnelle

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $\mathcal{G}$ une sous tribu de $\mathcal{F}$.

***Théorème 3.2* (Définition et Théorème).** — *Soit X une variable aléatoire positive ou intégrable. Alors, il existe une unique, presque sûrement, variable aléatoire notée $\mathbb{E}[X \mid \mathcal{G}]$ telle que*

1. *$\omega \to \mathbb{E}[X \mid \mathcal{G}](\omega)$ est $\mathcal{G}$ mesurable.*
2. *pour tout $B \in \mathcal{G}$:*

$$\mathbb{E}[X1_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]1_B]. \tag{19}$$

En particulier, on a :

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]]. \tag{20}$$

Démonstration. — **existence** Supposons d'abord X positive. La mesure positive

$$\mu(A) = \mathbb{E}[X1_A] \qquad (A \in \mathcal{F}), \tag{21}$$

est par construction absolument continue par rapport à $\mathbb{P}$ sur la tribu $\mathcal{F}$ (si $\mathbb{P}(A) = 0$, alors $\mu(A) = 0$). Sa restriction à la sous tribu $\mathcal{G}$ définit donc également une mesure absolument continue par rapport à $\mathbb{P}$, sur $\mathcal{G}$. On peut donc considérer sa dérivée de Radon-Nykodim, $Z = \frac{d\mu}{d\mathbb{P}_{|\mathcal{G}}}$ qui est par construction une variable aléatoire positive $\mathcal{G}$ mesurable telle que

$$\mu(B) = \mathbb{E}[X1_B] = \int Z1_B \, d\mathbb{P} = \mathbb{E}[Z1_B] \qquad (B \in \mathcal{G}). \tag{22}$$

Maintenant si X est une variable aléatoire intégrable, alors on décompose $X(\omega)$ en parties positive et négative : $X = X^+ - X^-$. On définit les espérances conditionnelles $\mathbb{E}[X^\pm \mid \mathcal{G}]$ qui sont bien des variables aléatoires $\mathcal{G}$ mesurables, positives et intégrables. Il est immédiat que $Z = \mathbb{E}[X^+ \mid \mathcal{G}] - \mathbb{E}[X^- \mid \mathcal{G}]$ satisfait les hypothèses 1 et 2.

unicité Si Y et Z conviennent alors $B = \{Y \geq Z\}$ est dans $\mathcal{G}$,donc :

$$\mathbb{E}[X1_B] = \mathbb{E}[Y1_B] = \mathbb{E}[Z1_B] \tag{23}$$

ce qui entraîne $0 = \mathbb{E}[(Y-Z)1_B]$ et donc $(Y-Z)\,\mathbf{1}_{(Y\geq Z)} = 0$ presque sûrement. De même, $(Z-Y)\,\mathbf{1}_{(Z\geq Y)} = 0$ presque sûrement, et donc $Y = Z$ presque sûrement.
En fait on peut avoir $\mathbb{E}[X1_B] = +\infty$ ce qui rend la preuve ci dessus caduque.Il suffit de remplacer B par $\{A \geq Y \geq Z\}$ avec $A > 0$ une constante, pour obtenir que

$$\mathbb{P}(A \geq Y, Y < Z) = 0\,.$$

On prend la réunion pour A dans $\mathbb{N}$ pour obtenir que

$$\mathbb{P}(Y < Z) = 0\,.$$

Par symétrie $\mathbb{P}(Z < Y) = 0$ ce qui termine la démonstration. □

Remarque. — *Lorsque X est intégrable,la condition 2 peut être remplacée par la condition suivante : pour tout variable aléatoire bornée $\mathcal{G}$ mesurable U,*

$$\mathbb{E}[XU] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]U]\,. \tag{2'}$$

Lorsque X est simplement positive, on demande juste à U d'être positive et $\mathcal{G}$ mesurable.
Il est immédiat de vérifier que lorsque $\mathcal{G}$ est engendrée par un système complet d'évènements, alors cette définition coïncide avec la définition (15).

Définition 3.3. — *On pose, pour X positive où intégrable,*

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)]\,. \tag{24}$$

D'après le Lemme de Doob, c'est donc une fonction mesurable de Y, $\mathbb{E}[X \mid Y] = \phi(Y)$ caractérisée par

$$\mathbb{E}[Xf(Y)] = \mathbb{E}[\phi(Y)g(Y)] \tag{25}$$

pour toute fonction mesurable g, bornée si X est intégrable, positive si X est positive.

Proposition 3.3. — *On se place sous les conditions d'application du Théorème 3.2 pour définir correctement toutes les espérances conditionnelles ci-dessous pour les variables aléatoires X, Y, X_n, ... Les égalités et inégalités ne sont valables que presque sûrement.*

1. *(linéarité)* $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$
2. *(monotonie) si $X \leq Y$, alors $\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}]$.*
3. *(convergence monotone) si $0 \leq X_n$ et $X_n \uparrow X$ presque sûrement, alors $\mathbb{E}[X_n \mid \mathcal{G}] \uparrow \mathbb{E}[X \mid \mathcal{G}]$.*
4. *(convergence dominée) si $X_n \to X$ presque sûrement et si $\sup_n |X_n| \leq Y$ et Y est intégrable, alors $\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}]$ presque sûrement.*

5. *(Jensen conditionnelle) Si ϕ est convexe et si $\phi(X)$ est positive ou intégrable,alors*

$$\phi(\mathbb{E}\left[X \mid \mathcal{G}\right]) \leq \mathbb{E}\left[\phi(X) \mid \mathcal{G}\right]. \tag{26}$$

   En particulier, si X est intégrable

$$|\mathbb{E}\left[X \mid \mathcal{G}\right]| \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right]$$

6. *(projection) Si $\mathcal{H} \subset \mathcal{G}$ est une sous tribu de $\mathcal{G}$ alors,*

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mid \mathcal{H}\right] = \mathbb{E}\left[X \mid \mathcal{H}\right].$$

7. *si X est indépendante de $\mathcal{G}$ alors $\mathbb{E}\left[X \mid \mathcal{G}\right] = \mathbb{E}\left[X\right]$ (l'espérance conditionnelle est constante).*
8. *si Y est $\mathcal{G}$ mesurable, alors*

$$\mathbb{E}\left[XY \mid \mathcal{G}\right] = Y\mathbb{E}\left[X \mid \mathcal{G}\right].$$

Démonstration. — Les propriétés 1 et 2 sont triviales. Pour la convergence monotone conditionnelle, soit Z la limite croissante de $\mathbb{E}\left[X_n \mid \mathcal{G}\right]$. Alors Z est positive $\mathcal{G}$ mesurable et pour tout $B \in \mathcal{G}$, par convergence monotone $\mathbb{E}\left[X_n 1_B\right] = \mathbb{E}\left[\mathbb{E}\left[X_n \mid \mathcal{G}\right] 1_B\right]$ converge vers $\mathbb{E}\left[X 1_B\right]$ et vers $\mathbb{E}\left[Z 1_B\right]$. Donc $Z = \mathbb{E}\left[X \mid \mathcal{G}\right]$. La convergence dominée conditionnelle se démontre comme la convergence dominée classique en utilisant un lemme de Fatou conditionnel !
Pour l'inégalité de Jensen conditionnelle, supposons par exemple que X et $\phi(X)$ soient intégrables, et posons $Y = \mathbb{E}\left[X \mid \mathcal{G}\right]$, $Z = \mathbb{E}\left[\phi(X) \mid \mathcal{G}\right]$. Alors, pour tout variable aléatoire U bornée $\mathcal{G}$ mesurable, et toute fonction linéaire $\psi(x) = ax + b$ qui minore ϕ, on a:

$$\begin{aligned}\mathbb{E}\left[ZU\right] = \mathbb{E}\left[\phi(X)U\right] &\geq \mathbb{E}\left[\psi(X)U\right] = \mathbb{E}\left[(aX+b)U\right] \\ &= \mathbb{E}\left[(aY+b)U\right] = \mathbb{E}\left[\psi(Y)U\right].\end{aligned}$$

En conséquence, en prenant $U = \mathbf{1}_{(\psi(Y) \geq Z)}$ on obtient

$$\mathbb{E}\left[(\psi(Y) - Z)\,\mathbf{1}_{(\psi(Y) \geq Z)}\right] \leq 0$$

et dons $Z \geq \psi(Y)$ presque sûrement. On note Ω_ψ un ensemble de probabilité 1, $\mathbb{P}\left(\Omega_\psi\right) = 1$, tel que pour tout $\omega \in \Omega_\psi$ on ait $Z(\omega) \geq \psi(Y(\omega))$.
Observons que $\phi(x) = \sup\left\{\psi(x) : \psi \in \Lambda\right\}$ avec Λ l'ensemble des fonctions $\psi(x) = ax + b$ linéaires, qui minorent ϕ et à coefficients a, b rationnels. En conséquence si $\Omega' = \cap_{\psi \in \Lambda} \Omega_\psi$, on a $\mathbb{P}\left(\Omega'\right) = 1$ et pour ω dans Ω',$Z(\omega) \geq \phi(Y(\omega))$. C'est à dire $Z \geq \phi(Y)$ presque sûrement.
6. On pose $Y = \mathbb{E}\left[X \mid \mathcal{G}\right]$ et $Z = \mathbb{E}\left[Y \mid \mathcal{H}\right]$. Soit $B \in \mathcal{H}$. Comme $B \in \mathcal{H}$, on a:

$$\mathbb{E}\left[X 1_B\right] = \mathbb{E}\left[Y 1_B\right] = \mathbb{E}\left[Z 1_B\right]$$

d'où le résultat.

7. Si $B \in \mathcal{G}$ et $\mu = \mathbb{E}[X]$, alors, par indépendance de X et 1_B,

$$\mathbb{E}[X 1_B] = \mathbb{E}[X]\mathbb{E}[1_B] = \mu\mathbb{E}[1_B] = \mathbb{E}[\mu 1_B]$$

Comme la variable constante égale à μ est bien $\mathcal{G}$ mesurable, on a bien $\mathbb{E}[X \mid \mathcal{G}] = \mu$.
8. Supposons X et Y positives. Si $B \in \mathcal{G}$, alors, $Y 1_B$ est $\mathcal{G}$ mesurable positive,donc

$$\mathbb{E}[XY 1_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}](Y 1_B)] = \mathbb{E}[(\mathbb{E}[X \mid \mathcal{G}]Y) 1_B],$$

d'où le résultat.
Si X et XYsont intégrables, on décompose X et Y en parties positive et négative et on applique la linéarité. □

Rappelons que $L^2(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de Hilbert, muni du produit scalaire $(X, Y) \to \mathbb{E}[XY]$ et que l'on peut voir $L^2(\Omega, \mathcal{G}, \mathbb{P})$ comme un sous espace fermé formé des variables de carré intégrable, qui sont $\mathcal{G}$ mesurables.

Proposition 3.4. — *Si X est de carré intégrable, alors $\mathbb{E}[X \mid \mathcal{G}]$ est la projection orthogonale de X sur le sous espace fermé $L^2(\Omega, \mathcal{G}, \mathbb{P})$.*

Démonstration. — En effet, si Z est cette projection orthogonale et si U est $\mathcal{G}$ mesurable bornée, alors $U \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ et donc U est orthogonale à $X - Z$, $0 = \mathbb{E}[U(X - Z)]$, ce qui entraîne $\mathbb{E}[XU] = \mathbb{E}[ZU]$. Comme Z est bien $\mathcal{G}$ mesurable, c'est l'espérance conditionnelle $\mathbb{E}[X \mid \mathcal{G}]$. □

SECTION 3

Conditionnement par une variable aléatoire

Soit Y une variable aléatoire: $Y : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$. La tribu engendrée par Y est

$$\sigma(Y) = Y^{-1}(\mathcal{E}) = \left\{Y^{-1}(B), B \in \mathcal{E}\right\}. \tag{27}$$

Soit X une autre variable aléatoire, positive ou intégrable, définie sur le même espace probabilisé $\Omega, \mathcal{A}, \mathbb{P})$. On note

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)]. \tag{28}$$

C'est une variable aléatoire réelle $\sigma(Y)$ mesurable, donc d'après le Lemme de Doob, il existe une fonction ϕ mesurable telle que

$$\mathbb{E}[X \mid Y] = \phi(Y) \qquad p.s. \tag{29}$$

Proposition 3.5. — *Si Y est une variable aléatoire discrète, et X une va positive ou intégrable, alors*

$$\mathbb{E}[X \mid Y] = \sum_y \mathbb{E}[X \mid Y = y] \mathbf{1}_{(Y(\omega)=y)} . \tag{30}$$

Démonstration. — Bien sûr la somme est sur les y, finis ou débombrables, tels que $\mathbb{P}(Y = y) > 0$. Il suffit de vérifier que la fonction $\phi(y) = \sum_y \mathbb{E}[X \mid Y = y]$ vérife pour tout $A \in \sigma(Y)$,

$$\mathbb{E}[X 1_A] = \mathbb{E}[\phi(Y) 1_A] . \tag{31}$$

Ceci est évident, par interversion somme intégrale, triviale dans le cas $X \geq 0$, car $1_A = \sum_{y \in A} 1_{\{y\}}$. Dans le cas X intégrable, on écrit $X = X^+ - X^-$ ou on utilise la convergence dominée pour les séries de fonctions. □

Exemple 3.2. — *Soit X le résulatt d'un dé équilibré et Y la variable aléatoire qui vaut i si X est impair, et p si X est pair. Déterminer $\mathbb{E}[X \mid Y]$.*
X suit la loi uniforme sur $\{1, \ldots, 6\}$ et $Y = f(X)$ avec $f(2) = f(4) = f(6) = p$ et $f(1) = f(3) = f(5) = i$. On a $\mathbb{P}(Y = i) = \mathbb{P}(Y = p) = \frac{1}{2}$ et

$$\mathbb{E}\left[X \mathbf{1}_{(Y=p)}\right] = \mathbb{E}\left[X \mathbf{1}_{(X=2)}\right] + \mathbb{E}\left[X \mathbf{1}_{(X=4)}\right] + \mathbb{E}\left[X \mathbf{1}_{(X=6)}\right] = \frac{1}{6}(2+4+6) = 2 . \tag{32}$$

De même, $\mathbb{E}\left[X \mathbf{1}_{(Y=i)}\right] = \frac{1}{6}(1+3+5) = 3/2$ et donc

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid Y = p] \mathbf{1}_{(Y=p)} + \mathbb{E}[X \mid Y = i] \mathbf{1}_{(Y=i)} = 3 \mathbf{1}_{(Y=i)} + 4 \mathbf{1}_{(Y=p)} . \tag{33}$$

Exemple 3.3. — *Soit $X \sim \mathcal{E}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ indépendantes. Alors, $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$ est une constante.*
En revanche, si $N = \mathbf{1}_{(X>Y)} + 2 \mathbf{1}_{(X \leq Y)}$ alors

$$\mathbb{E}[N \mid Y] = 2 - e^{-\lambda Y} \tag{34}$$

car, par indépendance,

$$\begin{aligned} \mathbb{E}\left[N \mathbf{1}_{(Y=k)}\right] &= \mathbb{P}(X > k, Y = k) + 2\mathbb{P}(X \leq k, Y = k) \\ &= \mathbb{P}(Y = k)(\mathbb{P}(X > k) + 2\mathbb{P}(X \leq k)) = \mathbb{P}(Y = k)(2 - e^{-\lambda k}) . \end{aligned}$$

On en déduit,

$$\mathbb{E}[N] = \mathbb{E}[\mathbb{E}[N \mid Y]] = 2 - \mathbb{E}\left[e^{-\lambda Y}\right] = 2 - e^{-\mu(1-e^{-\lambda})} . \tag{35}$$

Même lorsque la variable aléatoire Y n'est pas discrète, on a la caratérisation suivante:

Proposition 3.6. — *Si X est une variable positive (resp. intégrable), alors* $\mathbb{E}[X \mid Y] = \phi(Y)$, *$\phi$ mesurable positive (resp. $\phi(Y)$ intégrable) telle que*

$$\mathbb{E}[X\psi(Y)] = \mathbb{E}[\phi(Y)\psi(Y)], \tag{36}$$

pour toute fonction ψ mesurable positive (resp. bornée).

SECTION 4

Loi conditionnelle

Si on n'avait pas des identités presque sûres, mais de vraies identités, alors à ω fixé, l'application $f \to \mathbb{E}[f(X) \mid \mathcal{G}]$ aurait toutes les propriétés d'une mesure de probabilité. Il nous faut un argument topologique supplémentaire pour rendre cela rigoureux, et la définition d'un noyau.

Définition 3.4. — *Étant donnés deux espaces mesurés $(\Omega, \mathcal{F})$ et $(E, \mathcal{E})$, un* noyau de transition *de $(\Omega, \mathcal{F})$ sur $(E, \mathcal{E})$ est une application $q : \Omega \times \mathcal{E} \to [0, 1]$ qui vérifie :*

1. *pour tout $A \in \mathcal{E}$, la fonction $\omega \to q(\omega, A) = q_\omega(A)$ est $\mathcal{F}$ mesurable ;*
2. *pour tout $\omega \in \Omega$, la fonction $A \to q(\omega, A) = q_\omega(A)$ est une probabilité sur $(E, \mathcal{E})$.*

Pour toute fonction $f : (E, \mathcal{E}) \to \mathbb{R}$ borélienne positive ou bornée,on notera

$$\int f(x) dq_\omega(x) = \int f(x) q(\omega, dx)$$

la variable aléatoire espérance de f par rapport à la probabilité dépendant de ω.

Il est immédiat de vérifier que si $q : (\Omega, \mathcal{F}) \to \mathcal{M}_1(E)$ est une application mesurable, avec $\mathcal{M}_1(E)$ espace des mesures de probabilité sur E métrique séparable, muni de la topologie de la convergence faible, alors q induit un noyau de $(\Omega, \mathcal{F})$ sur $(E, \mathcal{E})$, par $q(\omega, A) = q(\omega)(A)$.

Théorème 3.7. — *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ une variable aléatoire à valeurs dans un espace polonais E muni de sa tribu des boréliens et soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$. Alors il existe une unique application mesurable $q : (\Omega, \mathcal{F}) \to \mathcal{M}_1(E)$ telle que*

$$\mathbb{E}[f(X) \mid \mathcal{G}] = \int f(x) q(\omega, dx)$$

pour tout fonction $f : (E, \mathcal{E}) \to \mathbb{R}$ borélienne positive ou bornée. Ce noyau est appelé loi conditionnelle régulière *de X étant donné $\mathcal{G}$.*

Démonstration. — Voir Borkar [**?**].

□

Une conséquence de ce théorème est l'existence de lois conditionnelles pour les couples de variables aléatoires.

Proposition 3.8. — *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ une variable aléatoire à valeurs dans un espace polonais E muni de sa tribu des boréliens, et soit $Y(\Omega, \mathcal{F}, \mathbb{P}) \to (E', \mathcal{E}')$ une variable aléatoire à valeurs dans un espace métrique séparable. Alors il existe un noyau K de $(E', \mathcal{E}')$ sur $(E, \mathcal{E})$ tel que pour toute fonction $\phi : (E, \mathcal{E}) \to \mathbb{R}$ borélienne bornée ou positive,*

$$\mathbb{E}[\phi(X) \mid Y](\omega) = \int \phi(x) dK_{Y(\omega)}(x) \qquad ps. \tag{37}$$

La mesure $K_y(dx) = K(y, dx)$ est appelée loi conditionnelle de X sachant $Y = y$ et notée parfois $\mathcal{L}(X \mid Y = y)$ ou $P_X^{Y=y}$ ou $P_{X|Y=y}$.

Démonstration. — On applique le théorème précédent à la tribu $\mathcal{G} = \sigma(Y)$. Comme l'application q est $\sigma(Y)$ mesurable, et que Y est à valeurs dans un espace métrique séparable, le Lemme de Doob **??** assure l'existence d'une application mesurable $K : (E', \mathcal{E}') \to \mathcal{M}_1(E)$ telle que $q(\omega) = K(Y(\omega))$, d'où le résultat. □

Corollaire 3.9. — *Si X, Y sont des variables aléatoires à valeurs dans des espaces polonais, et ϕ est mesurable positive (ou intégrable)*

$$\begin{aligned}\mathbb{E}[\phi(X,Y)] &= \mathbb{E}\left[\int \phi(x,Y) K_Y(dx)\right] = \mathbb{E}\left[\mathbb{E}[\phi(X,y) \mid Y = y]_{|y=Y(\omega)}\right] \\ &= \int \mathbb{E}[\phi(X,y) \mid Y = y]\, dP_Y(y)\,.\end{aligned}$$

Le cadre d'application le plus courant est celui des variables aléatoires à valeurs dans $\mathbb{R}$ ou $\mathbb{R}^d$. L'existence d'une densité pour le couple permet de donner une forme explicite à la loi conditionnelle.

Proposition 3.10. — *Soit (X, Y) un couple de variables aléatoires réelles admettant la densité f. Alors la loi conditionnelle de X sachant $Y = y$, $\mathcal{L}(X \mid Y = y)$, admet pour densité pour $y \in A$,*

$$f_{Y=y}(x) = \frac{f(x,y)}{\int f(x',y)\, dx'} \qquad \text{avec} \quad A = \left\{y : \int f(x',y)\, dx' > 0\right\}. \tag{38}$$

Démonstration. — C'est une application du théorème de Fubini. On rappelle que par la méthode de la fonction muette, Y admet pour densité

$$f_Y(y) = \int f(x,y)\, dy\,. \tag{39}$$

En particulier,

$$\mathbb{P}\left(Y \in A^C\right) = \int 1_{A^C}(y) f_Y(y)\, dy = \int \mathbf{1}_{(f_Y(y)=0)} f_Y(y)\, dy = 0\,. \tag{40}$$

En conséquence, si ϕ est mesurable positive, par Fubini positif

$$\begin{aligned}
\mathbb{E}\left[\phi(X,Y)\right] = \mathbb{E}\left[\phi(X,Y)1_A(Y)\right] &= \int \phi(x,y) f(x,y) 1_A(y) dxdy \\
&= \int \left(\int \phi(x,y) f_{Y=y}(x)\right) f_Y(y) 1_A(y) dy \\
&= \mathbb{E}\left[\int \phi(x,Y) K_Y(dx)\right],
\end{aligned}$$

avec $K_y(dx) = f_{Y=y}(x)dx\, \mathbf{1}_{(y\in A)} + \nu(dx)\, \mathbf{1}_{(y\notin A)}$ avec ν une probabilité fixe (qui ne sert à rien à part à définir K_y pour les valeurs que y ne prend pas!).

$\square$

Remarque. — *On peut également appliquer la proposition précédente si on a l'existence d'une densité du couple par rapport au produit $\mu \otimes \nu$ de deux mesures σ finies.*

Exemple 3.4. — *Soit (T_1, T_2) IID de loi $\mathcal{E}(\lambda)$. Déterminer $\mathbb{E}\left[T_1 \mid T_1 + T_2\right]$. Première méthode : on peut évidemment calculer la loi du couple $(T_1, T_1 + T_2)$ par la méthode de la fonction muette et un changement de variable. Elle admet pour densité $f(x, y) = \lambda^2 e^{-\lambda y}\, \mathbf{1}_{(y>x>0)}$ donc $T_1 + T_2$ admet pour densité $f_Y(y) = \lambda^2 y e^{-\lambda y}\, \mathbf{1}_{(y>0)}$ (c'est la la loi $\gamma(2, \lambda)$) et pour $y > 0$, $f_{Y=y}(x) = \frac{1}{y}\, \mathbf{1}_{(y>x>0)}$, c'est la loi $\mathcal{U}(0, y)$ et donc*

$$\mathbb{E}\left[T_1 \mid T_1 + T_2 = y\right] = \mathbb{E}\left[\mathcal{U}(0, y)\right] = y/2\,. \tag{41}$$

Dexuxième méthode : on peut trouver plus simplement le résultat précédent en remarquant que $\mathbb{E}\left[T_1 \mid T_1 + T_1\right] = h(T_1 + T_2)$ avec h mesurable positive. Comme (T_1, T_2) et (T_2, T_1) ont même loi $\mathcal{E}(\lambda) \otimes \mathcal{E}(\lambda)$, $(T_1, T_1 + T_2) \overset{d}{=} (T_2, T_1 + T_2)$ et donc si ϕ est mesurable positive

$$\mathbb{E}\left[T_2 \phi(T_1 + T_2)\right] = \mathbb{E}\left[T_1 \phi(T_1 + T_2)\right] = \mathbb{E}\left[h(T_1 + T_2)\phi(T_1 + T_2)\right],$$

d'où l'on déduit que $\mathbb{E}\left[T_2 \mid T_1 + T_1\right] = h(T_1 + T_2)$ et donc par linéarité

$$T_1 + T_2 = \mathbb{E}\left[T_1 + T_2 \mid T_1 + T_2\right] = \mathbb{E}\left[T_1 \mid T_1 + T_1\right] + \mathbb{E}\left[T_2 \mid T_1 + T_1\right] = 2h(T_1 + T_2)\,. \tag{42}$$

En définitive, $\mathbb{E}\left[T_1 \mid T_1 + T_2\right] = \frac{1}{2}(T_1 + T_2)$.

On retient parfois l'existence d'une loi conditionnelle régulière sous la forme de la *formule de désintégration*

$$
\boxed{dP_{X,Y}(x,y) = dP_{X|Y=y}(x)\, dP_Y(y)}\,, \tag{43}
$$

que l'on peut utiliser dans les deux sens : si on connait la loi du couple, on peut en déduire la loi marginal et la loi conditionnelle. Si on connaît la loi marginale, et la loi conditionnelle, alors on connait la loi du couple. En particulier, ce sont des lois absolument continues en même temps.

SECTION 5

Conditionnement et indépendance

On se donne Y et Z deux variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans des espaces polonais F et G, et $f : F \times G \to E$ une application mesurable à valeurs dans un espace polonais E.

Proposition 3.11. — *Soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$ telle que Y soit $\mathcal{G}$ mesurable et Z soit indépendante de $\mathcal{G}$. Alors la loi conditionnelle régulière de $X = f(Y,Z)$ sachant $\mathcal{G}$ est donnée par (ϕ étant borélienne positive ou bornée)*

$$
\mathbb{E}\left[\phi(X) \mid \mathcal{G}\right] = \int \phi(f(Y(\omega), z))\, dP_Z(z) = \mathbb{E}\left[\phi(f(y,Z))\right]_{|y=Y(\omega)} \tag{44}
$$

En particulier, la loi conditionnelle régulière de $X = f(Y,Z)$ sachant $Y = y$ est la loi de la variable aléatoire $f(y,Z)$.

Démonstration. — Le cas particulier s'obtient en prenant $\mathcal{G} = \sigma(Y)$. Pour le cas général soit $\phi : E \to \mathbb{R}$ borélienne, disons bornée. Soit $B \in \mathcal{G}$. Alors, l'espace $\mathcal{H}$ des fonctions $g : F \times G \to \mathbb{R}$ boréliennes bornées telles que

$$
\mathbb{E}\left[g(Y,Z)1_B\right] = \mathbb{E}\left[\left(\int g(Y,z)\, dP_Z(z)\right)1_B\right]
$$

est un espace vectoriel monotone, qui contient la classe $\mathcal{C}$ des fonctions de la forme $g(y,z) = h(y)k(z)$, h,k boréliennes bornées. En effet, d'après le théorème de Fubini-Tonelli, et l'indépendance de $k(Z)$ et $h(Y)1_B$, on a

$$
\begin{aligned}
\mathbb{E}\left[g(Y,Z)1_B\right] &= \mathbb{E}\left[h(Y)1_B k(Z)\right] = \mathbb{E}\left[h(Y)1_B\right]\mathbb{E}\left[k(Z)\right] \\
&= \mathbb{E}\left[h(Y)1_B\right] \int k(z)\, dP_Z(z) \\
&= \mathbb{E}\left[\left(\int h(Y)k(z)\, dP_Z(z)\right)1_B\right] \\
&= \mathbb{E}\left[\left(\int g(Y,z)\, dP_Z(z)\right)1_B\right].
\end{aligned}
$$

En vertu du théorème des classes monotones, $\mathcal{H}$ contient toutes les fonction $\sigma(\mathcal{C})$ bornées. Comme $\sigma(\mathcal{C})$ est la tribu des boréliens, on a

$$\mathbb{E}\left[\phi(f(Y,Z))1_B\right] = \mathbb{E}\left[\left(\int \phi(f(Y,z))\,dP_Z(z)\right)1_B\right]$$

ce qui prouve (44) puisque B est arbitraire.

$\square$

4

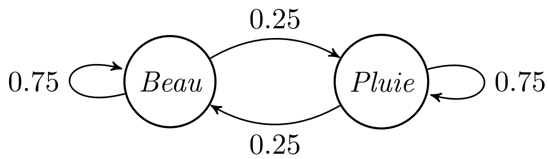
Chaîne de Markov à temps discret

SECTION 1 Introduction

Exemple 4.1. — *On dispose de statistiques, de tableaux de chiffres, indiquant qu'il fait beau 50% du temps et mauvais 50% du temps. Une première façon naïve de prédire le temps est la suivante*

$$\mathbb{P}\,(\textit{beau}) = \mathbb{P}\,(\textit{mauvais}) = 0.5$$

La justification est la loi des grands nombres : l'hypothèse de base est que le climat du lendemain est indépendant du climat des autres jours.
Malheureusement, on constate en examinant les chiffres qu'il y a 3 fois plus de chances que le climat du lendemain reste le même que celui d'aujourd'hui (plutôt qu'il ne change). On représente ce modèle par le diagramme

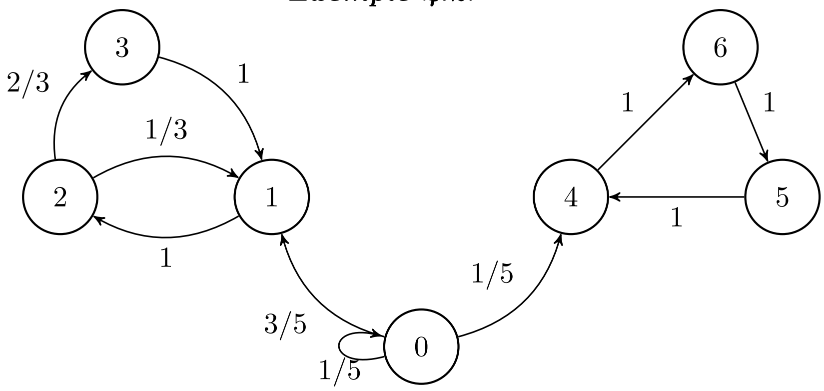


ou encore par la matrice $P = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$.
Pour prédire le temps de demain nous avons besoin de

- *un mécanisme de transition (décrit par un graphe ou une matrice)*
- *savoir quel temps il fait aujourd'hui*

Considérons un exemple plus évolué

Exemple 4.2. —



On aimerait répondre à la liste de questions suivantes:

a) Partant de 0 *quelle est la probabilité de toucher 6 ?*
b) Partant de 1 quelle est la probabilité de toucher 3 ?
c) partent de 1 combien de temps en moyenne faut-il pour toucher 3 . (réponse 3)
d) Partant de 1, la proportion de temps passé en 2 est, en temps long, 3/8
e) Partant de 0, la probabilité que je sois en 1 à l'instant n, est pour n grand, proche de $9/32$.

SECTION 2

Définitions et propriétés de base

Soit I un emsemble fini ou dénombrable, applelé espace d'états (nommé aussi S ou E). Une mesure (positive) sur I est une famille $(\lambda_i, i \in I)$ de nombres positifs. C'est une loi de probabilité si $\sum_i \lambda_i = 1$.
Une matrice $P = (p_{ij}$ est stochastique (on dit aussi matrice de transition) si toutes ses lignes sont des probabilités i.e.

$$\forall i, j \, p_{ij} \geq 0 \quad \text{et} \quad \forall i, \sum_j p_{ij} = 1$$

Il y a une bijection évidente entre les matrices de transition et les garphes valués donnée par il y a une flèche de i à j valuée par p_{ij} ssi $p_{ij} > 0$.

Définition 4.1. — *Une chaîne de Markov homogène à valeurs dans I de matrice de transition P et de loi initiale λ est une famille de variables aléatoires* $(X_n, n \in \mathbb{N})$ *telle que*

1. $\mathbb{P}(X_0 = i) = \lambda(i)$
2. $\mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n) = p_{i_n, j_n}$

On dit alors que $(X_n)_{n\in\mathbb{N}}$ est markov (λ, P).

Exemple 4.3. — *$I = \{1, 2, 3\}$ est l'espace des 'états avec 1 qui signifie en panne, 2 en fonctionnement erratique et en bon fonctionnement. Les transitions sont décrites par la matrice*

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

Question *Si je pars d'une machine en bon état, quelle est la probabilité α (resp. β) qu'elle soit en bon état (resp. erratique) dans 4 jours ?*
Réponse *il y a deux chemins de même probabilité pour obtenir une machine en bon état $\alpha = 2\frac{1}{2}^2\frac{1}{3}^2$*

2.1 Caractérisation

Proposition 4.1. — *$(X_n)_{0\leq n\leq N}$ est Markov λ, P ssi pour tous $i_0, ..., i_N$*

$$\mathbb{P}\left(X_0 = i_0, \ldots, X_N = i_N\right) = \lambda(i_0)p_{i_0,i_1}...p_{i_{N-1},i_N}$$

Démonstration. — conditionnements successifs et récurrence. □

On note δ_i la masse de Dirac en i et $\mathbb{P}_i$ la loi d'une chaîne (δ_i, P) (on dit que la chaîne est issue de i car $X_o = i$ ps).

Proposition 4.2. — *Si $(X_n)_{n\geq 0}$ est Markov (λ, P), alors conditionnellement à $X_m = i$ le processus $(X_{m+n}, n \geq 0)$ est Markov (δ_i, P) et est indépendant de $(X_0, \ldots, X_m)$.*

Démonstration. — Il suffit de montrer que pour tout évènement $A \in \mathcal{F}_m^X := \sigma(X_0, ..., X_m)$ et tout évènement $B = \{X_m = i_m, \ldots, X_{m+n} = i_{m+n}\}$ de $\sigma(X_p, p \geq m)$ on a

$$\mathbb{P}\left(A \cap B \mid X_m = i\right) = \mathbb{P}\left(A \mid X_m = i\right)\mathbb{P}_i(X_0 = i_m, \ldots, X_n = i_{m+n}) \tag{45}$$

En effet les évènements B de cette forme constituent un π système qui engendre $\sigma(X_p, p \geq m)$. On démontre (45) en prenant A de la forme $\{X_0 = i_0, \ldots, X_m = i_m\}$ car tout évènement de $\mathcal{F}_m^X$ est une réunion disjointe de tels ensembles. □

Etant donné une variable aléatoire $Z \in \mathcal{F}_\infty^X$, par le Lemme de Doob, $Z = F(X_n, n \geq 0)$ avec F une fonction mesurable,et donc on peut considèrer l'opérateur de shift $Z \circ \theta^m = F(X_{m+n}, n \geq 0)$ (et on note $\theta = \theta^1$). Alors, la proposition précédente s'écrit

Proposition 4.3. — *Pour tout variable aléatoire Z, $\mathcal{F}_\infty^X$ mesurable, positive ou bornée, et tout $m \geq 0$,*

$$\mathbb{E}\left[Z \circ \theta^m \mid \mathcal{F}_m\right] = \mathbb{E}_{X_m}\left[Z\right] \quad p.s.$$

Remarque. — *En appliquant la proposition précédente à* $m = 0$ *on obtient que quelle que soit la loi initiale* λ *de* X_0*, du moment qu'elle charge le point* i*,* ($\lambda(i) > 0$)*, on a pour tout évènement* A *,* $\mathbb{P}_i(A) = \mathbb{P}(A \mid X_0 = i)$*. En bref, conditionner la chaîne à valoir* i *en 0 revient à la faire partir de* i*.*

2.2 Calculs algébriques

On note une mesure positive λ come un vecteur ligne et donc λP est la mesure $(\lambda P)(j) = \sum_i \lambda(i) p_{i,j}$ On note $p_{i,j}^{(n)}$ l'élément i, j de la matrice P^n et donc on a $p_{i,j}^{(2)} = \sum_k p_{i,k} p_{k,j}$ et en général

$$p_{i;j}^{(n)} = \sum_{i_0 = i, i_1, \ldots, i_n = j} p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}$$

Proposition 4.4. — *Soit* $(X_n)_{n \geq 0}$ *Markov* (λ, P)*. ALors pour toutn la loi de* X_n *est* λP^n *i.e.*

$$\mathbb{P}(X_n = j) = (\lambda P^n)(j)$$

En particulier, pour $\lambda = \delta_i$ *et tout* $m \geq 0$*, on obtient*

$$\mathbb{P}_i(X_n = j) = \mathbb{P}(X_{m+n} = j \mid X_m = i) = p_{i,j}^{(n)}$$

Démonstration. — Formule des probabilités totales :

$$\mathbb{P}(X_n = j) = \sum_{i_0, \ldots, i_n = j} \mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)$$

□

Exemple 4.4. — *Pour le climat à Nantes, s'il fait beau aujourd'hui, la probabilité qu'il fasse beau dans trois jours est*

$$\mathbb{P}(X_3 = 1 \mid X_0 = 1) = p_{i,j}^{(3)} = 9/16$$

car on calcule

$$P^2 = \begin{pmatrix} 10/16 & 6/16 \\ 6/16 & 10/16 \end{pmatrix}$$

2.3 Autres propriétés

Définition 4.2. — *On appelle filtration une suite croissante de tribus* $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ *sur* Ω*. On dit que la va* $T : \omega \to \bar{N}$ *est un temps d'arrêt si pour tout* n*,* $\{T \leq n\} \in \mathcal{F}_n$*. On dit qu'un processus* Z_n *est adapté si pour tout* n *la va* Z_n *est* $\mathcal{F}_n$ *mesurable.*

Pour une chaîne de Markov $(X_n, n \in \mathbb{N})$ on considère la filtration naturelle $\mathcal{F}_n = \sigma(X_k, k \leq n)$ et $\mathcal{F}_\infty = \sigma X_n, n \geq 0$. Un des temps d'arrêt le plus utilisé est le premier temps de retour en i

$$T_i = \inf \{n \geq 1 : X_n = i\}$$

avec la convention $\inf \emptyset = +\infty$. On note $\mathcal{F}_T = \sigma\{Z_T, A \text{ processus adapté}\}$ la filtration arrêtée en T.

Proposition 4.5. — *Pour toute va Z $\mathcal{F}_\infty$ mesurable, positive ou bornée, et tout temps d'arrêt T, sur $T < +\infty$ on a p.s.*

$$\mathbb{E}\left[Z \circ \theta^T \mid \mathcal{F}_T\right] = \mathbb{E}_{X_T}\left[Z\right]$$

SECTION 3

Classification des états

3.1 Structure de classe

Définition 4.3. — *On dit que i mène à j et on le note $i \to j$ si*

$$\mathbb{P}_i(\exists n \geq 0 : X_n = j) > 0.$$

On dit que i et j communiquent et on le note $i \leftrightarrow j$ si $i \to j$ et $j \to i$.

Proposition 4.6. — *Si $i \neq q$ alors sont équivalentes*

(a) $i \to j$
(b) $\exists i_0 = i, i_1, \ldots, i_n = j$ tel que $p_{i_0,i_1} \cdots p_{i_{n-1},i_n} > 0$
(c) Il existe un chemin dans le graphe de transition menant de i à j.
(d) $\exists n \geq 0$ tel que $p_{i,j}^{(n)} > 0$.

Démonstration. — (b) et (c) sont identiques; Comme une somme de termes positifs est > 0 ssi il existe un term > 0, on en déduit que $(b) \iff (d)$. Enfin l'équivalence entre (a) et (d) découle de la double inégalité

$$p_{i,j}^{(n)} = \mathbb{P}_i(X_n = j) \leq \mathbb{P}_i(\exists m, X_m = j) \leq \sum_m \mathbb{P}_i(X_m = j)$$

En effet, si (a), alors $\mathbb{P}_i(\exists m, X_m = j) > 0$ donc $\sum_m \mathbb{P}_i(X_m = j) > 0$ et il existe m tel que $\mathbb{P}_i(X_m = j)$. □

Corollaire 4.7. — *la relation $i \leftrightarrow j$ est une relation d'équivalence. Les classes d'équivalence sont appélées classes de communication.*

Démonstration. — Pour la transitivité si $i \to j$ et $j \to k$ alors il existe m, n tels que $p_{i,j}^{(n)} > 0$ et $p_{j,k}^{(m)} > 0$ donc $p_{i,k}^{(n+m} \geq p_{i,j}^{(n)} p_{j,k}^{(m)} > 0$. □

Définition 4.4. — *La chaîne est dite irréductible s'il y a une seule classe. La classe C est dite ouverte s'il existe $i \in C$ et $j \notin C$ tels que $i \to j$.*

Proposition 4.8. — *La classe C est ouverte ssi il existe $i \in C$ et $j \notin C$ tels que $p_{i,j} > 0$.*
On suppose la classe fermée et que p.s. $X_0 \in C$. Alors presque sûrement $\forall n, X_n \in C$. En d'autres termes une chaîne reste toujours p.s. dans une classe fermée.

Démonstration. — Pour la première partie il existe un chemin fini de C à l'extérieur de C. Il y a bien un moment ou il quitte C.
Pour la seconde partie on fait une récurrence et on voit que l'hérédité découle de la remarque suivante. cOMME C est fermée, si $i \in C$, alors pour tout $j \notin C$, $p_{i,j}^{(n)} = 0$ et donc en sommant sur tous les $j \notin C$, $\mathbb{P}_i\big(X_n \in C^C\big) = 0$ et donc

$$\mathbb{P}\left(X_n \notin C\right) = \sum_i \mathbb{P}\left(X_0 = i\right)\mathbb{P}_i(X_n \notin C) = 0$$

□

3.2 Récurrence et transience

Définition 4.5. — *On dit que i est transient si $\sum_n p_{ii}^{(n)} < +\infty$ et récurrent si $\sum_n p_{ii}^{(n)} = +\infty$.*

Proposition 4.9. — *La récurrence est une proriété de classe. Soit C une classe de communication. ALors soit tous les éléments de C sont récurrents, soit ils sont tous transients.*

Démonstration. — soit $i, j \in C$ on a $i \leftrightarrow j$ donc il existe k, m tels que $p_{ij}^{(m)} > 0$ et $p_{ji}^{(k)} > 0$ On déduit de l'inégalité

$$p_{ii}^{(n+m+k)} \geq p_{ij}^{(m)} p_{jj}^{(n)} p_{ji}^{(k)}$$

que les séries $\sum_n p_{ii}^{(n)}$ et $\sum_n p_{jj}^{(n)}$ sont de même nature. □

Le nombre de retours en i (ou temps passé en i à partir de 1) est $N_i = \sum_{n\geq 1} \mathbf{1}_{(X_n=i)}$ et la probabilité de retour en i est $f_i := \mathbb{P}_i(N_i > 0) = \mathbb{P}_i(T < +\infty)$ avec $T = \inf\{n \geq 1 : X_n = i\}$.

Lemme 4.10. — *pour tout $r \in \mathbb{N}$, $\mathbb{P}_i(N_i > r) = f_i^{r+1}$.*

Démonstration. — SOit $T_0 = 0$ et $T_{k+1} = \inf n > T_k : X_n = 1$ le $k+1$ ième temps de passage en i (avec la convention $\inf \emptyset = +\infty$. Alors $\{N_i > r\} =$

$\{T_r < +\infty\}$ et on a $T_{k+1} = T_k + T_1 \circ \theta^{T_k}$, donc par la propriété de Markov forte,comme $X_{T_k} = i$ lorsque $T_k < +\infty$,

$$\begin{aligned}\mathbb{P}_i(N_i > k+1) = \mathbb{P}_i\big(T_k + T_1 \circ \theta^{T_k} < +infty\big) &= \mathbb{P}\left(T_k < +\infty, T_1 \circ \theta^{T_k} < +infty\right) \\ &= \mathbb{E}_i\left[\mathbf{1}_{(T_k<+\infty)}\mathbb{E}_{X_{T_k}}\left[\mathbf{1}_{(T_1<+\infty)}\right]\right] \\ &= \mathbb{P}_i(T_1 < +\infty)\mathbb{P}_i(T_k < +\infty) = f_i f_i^r \,.\end{aligned}$$

□

***Théorème 4.11* (dichotomie).** — *Si $f_i = 1$ alors presque sûrement $X_n = i$ pour une infinité d'entiers n et i est un état récurrent.*
Si $f_i < 1$, alors presque sûrement $X_n = i$ pour seulement un nombre fini d'entiers et l'état i est transient.

Démonstration. — On observe que d'une part

$$\mathbb{E}_i\left[N_i\right] = \sum_{k \in N} \mathbb{P}_i(N_i > k) = \sum_k f_i^{k+1} = \begin{cases} \frac{f_i}{1-f_i} & \text{si } f_i < 1; \\ +\infty & \text{si } f_i = 1; \end{cases}$$

et d'autre part

$$\mathbb{E}_i\left[N_i\right] = \sum_n \mathbb{E}_i\left[\mathbf{1}_{(X_n=i)}\right] = \sum_n p_{ii}^{(n)}$$

d'où la dichotomie. Si $f_i = 1$ alors $N_i = +\infty$ ps (car pour tout k, $\mathbb{P}_i(N_i > k) = 1$) et i est transient, alors que si $f_i < 1$ on a $N_i < +\infty$ ps et i est récurrent. □

Remarque. — *On montre de la même façon que si C est une classe récurrente et $i, j \in C$ alors $\mathbb{P}_i$ ps, on a $N_j = +\infty$ et $\mathbb{E}_i\left[N_j\right] = \sum_n p_{ij}^{(n)} = +\infty$ Si C est une classe transiente et $i, j \in C$ alors $\mathbb{P}_i$ ps, on a $N_j < +\infty$ et $\mathbb{E}_i\left[N_j\right] = \sum_n p_{ij}^{(n)} < +\infty$*

Proposition 4.12. — *Soit C une classe de communication.*

1. *Si C est ouverte alors C est transiente.*
2. *Si C et ouverte et C est finie, alors p.s. il existe $n_0 = n_0(\omega)$ tel que $\forall n \geq n_0, X_n \notin C$ (au bout d'un moment on sort définitivement d'une classe ouverte finie).*
3. *Si C est fermée et C est finie alors C est récurrente.*

Démonstration. — 1) Soit $i \in C, j \notin C$ tels que $p_{ij} > 0$ On a $j \not\to i$ donc, $1 = \mathbb{P}_j(\forall n \geq 0, X_n \neq i)$. Donc, par Markov,

$$\mathbb{P}_i(N_i = 0) \geq \mathbb{P}_i(X_1 = j, \forall n \geq 1 X_n \neq i) = p_{ij}\mathbb{P}_j(\forall n \geq 0, X_n \neq i) = p_{ij} > 0$$

i.e. $f_i = \mathbb{P}_i(N_i > 0) < 1$ et i est transient.
2) Si C est finie, on note $N_C = \sum_{j \in C} N_j$ le temps passé en C, et on a $\mathbb{E}_i\left[N_C\right] = \sum_{j \in C} \mathbb{E}_i\left[N_j\right]$ qui est une somme finie. Donc si C est en outre ouverte, alors

elle est transiente et pour tout $j \in C$ on a $\mathbb{E}_i[N_j] < +\infty$ donc $\mathbb{E}_i[N_C] < +\infty$ et donc $N_C < +\infty$ ps.
3) Si C est finie et fermée, et que $i \in C$, alors $X_0 \in C$ ps, et on a vu que ps $\forall n, X_n \in C$ et donc $N_C = +\infty$ ps, donc $\mathbb{E}_i[N_c] = +\infty$ et il existe j tel que $\mathbb{E}_i[N_j] = +\infty$, donc i est récurrent. □

Corollaire 4.13. — *Une chaîne de Markov irréductible sur un espace d'états fini est récurrente.*

Remarque. — *Sur un espace d'étas inifini tout peut arriver. Sur $\mathbb{Z}$ la ma simple est récurrente (TCL) alors que la ma biaisée est transiente (LGN : $S_n \to +\infty$)*

SECTION 4

Temps d'atteinte et probabilités d'absorption

On considère une chaîne de Markov, à espace d'états fini, qui a des états transients et des états absorbants, et on met la matrice de transition sous la forme canonique en plaçant les états transients au début.

$$P = \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right)$$

Proposition 4.14. — *Pour une chaîne de Markov avec un état absorbant, presque sûrement la chaîne sera absorbée, i.e. $Q^n \to 0$*

Proposition 4.15. — *Pour une chaîne de Markov avec un état absorbant, la matrice $I - Q$ est inversible d'inverse $N = \sum Q^n$ et n_{ij} est le nombre moyen de fois où la chaîne est dans l'état j, si elle part de l'état i.*

N est la matrice fondamentale.

Démonstration. — On appelle A l'ensemble des états absorbants. Alors, si $i, j \notin A$ $n_{ij} = \mathbb{E}_i\left[\sum_k \mathbf{1}_{(X_k=j)}\right] = \sum_k q_{ij}^{(k)} = (I-Q)_{ij}^{(-1)}$.
□

Théorème 4.16. — *Soit t_i le temps moyen d'absorption de la chaîne si elle part dans l'état i. Alors, $t = N1$. En particulier, si $i \notin A$, alors $t_i = 1 + \sum_{j \notin A} p_{ij} t_j = 1 + (Qt)_i$.*

Démonstration. — $(N1)_i = \sum_j n_{ij}$ □

Théorème 4.17. — *Soit b_{ij} la probabilité que la chaîne partant de i soit absorbée en j. Alors, $B = NR$.*

Démonstration. — $B_{ij} = \sum_{n,k} \mathbb{P}\left(X_n = k, X_{n+1} = j \mid X_0 = i\right) = \sum_{nk} q_{ik}^{(n)} r_{kj} = (NR)_{ij}$.

□

Exemple 4.5. — *espace d'états* $\{1,2,3,4\}$, $p_{i,i+1} = 0.4 = 1 - p_{i,i-1}$ *si* $i = 2, 3$ *,* 1 *et* 4 *absorbants. Alors on réordonne les états* $E = \{2,3,1,4\}$ *et on calcule* $Q = \begin{pmatrix} 0 & 04 \\ 0.6 & 0 \end{pmatrix}$ *et* $R = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.4 \end{pmatrix}$. *On trouve* $N = \begin{pmatrix} 1.32 & 0.526 \\ 0.789 & 1.32 \end{pmatrix}$ *et* $B = NR = \begin{pmatrix} 0.789 & 0.21 \\ 0.473 & 0.52 \end{pmatrix}$. *Si* 1 *est gagnant et* 4 *perdant et que l'on me donne deux fois ma mise quand je pars de 3, et une seule fois quand je pars de 2, quelle doit être ma stratégie ? (on part de 3 car* $20.473 > 0.789$*).*

On a parfois besoin de connaître la loi du temps de sortie d'un point.

Lemme 4.18. — *On suppose que la chaîne est issue de* i *et que* i *est non absorbant. Soit* $T = \inf\{n \geq 1 : X_n \neq i\}$ *Alors sous* $\mathbb{P}_i$*, les variables aléatoires* X_T *et* T *sont indépendantes,* T *de loi géométrique de paramètre* $1 - p_{ii}$ *et* X_T *de loi*

$$\mathbb{P}\left(X_T = j\right) = \frac{p_{ij}}{1 - p_{ii}}$$

Exemple 4.6. — *On revient sur l'exeple 3; Partant de* 0 *la proba de toucher* 6 *est donc la proba* $\mathbb{P}_0(X_T = 4) = 1/4$ *et la probabilité d'aller dans la classe fermée* $\{1,2,3\}$ *est* $3/4$. *Le temps pour toucher* 3 *peut être* $+\infty$ *si je suis absorbé dans la classe fermée* $\{4,5,6\}$. *Donc la moyenne de ce temps est* $+\infty$.

De façon plus générale, si A est une prtie de l'espace d'états, non nécessairement absorbante, on définit

$$H_A = \inf\{n \geq 0 : X_n \in A\} \qquad (\inf \emptyset = +\infty) \tag{46}$$

et les probabilités d'abosorption en A:

$$h_A(i) = \mathbb{P}_i(H_A < +\infty)\,. \tag{47}$$

Une simple application de la propriété de Markov, à l'instant 1, montre que h_A est solution du système

$$h(i) = \begin{cases} 1 & \text{si } i \in A\,. \\ \sum_j p_{ij} h(j) & \text{sinon}\,. \end{cases} \tag{48}$$

On peut même montrer que h_A est la solution positive bornée minimale du système précédent.

SECTION 5

Mesures Invariantes

On dit que la mesure positive λ est invariante (ou encore stationnaire) si $\lambda P = \lambda$.

Proposition 4.19. — *Si $(X_n)_{n\in\mathbb{N}}$ est Markov (λ, P) et λ est invariante, alors la loi de X_n est constante et vaut λ. En outre pour tout m, $(X_{m+n}, n \geq 0)$ est Markov (λ, P)*

Démonstration. — En effet $X_n \sim \lambda P^n = \lambda$ et

$$\mathbb{P}\left((X_{m+n}, n \geq 0) \in A \mid \mathcal{F}_n\right) = \mathbb{P}\left(\{(X_n, n \geq 0) \in A\} \circ \theta^m\right) = \mathbb{P}_{X_m}(\{(X_n, n \geq 0) \in A\})$$

et on prend les espérances. □

Les limites des lois de X_n sont automatiquement des probabilités invariantes sur un espace fini.

Proposition 4.20. — *On suppose I fini et que pour un $i_0 \in I$ on ait*

$$\forall j, \quad p^{(n)}_{i_0,j} \to \pi_j$$

Alors π est une probabilité invariante.

Démonstration. — Par convergence dominée

$$\pi_j = \lim p^{(n)}_{i_0,j} = \lim p^{(n+1)}_{i_0,j} = \lim \sum_k p^{(n)}_{i_0,k} p_{kj} = \sum_k \pi_k p_{kj} = (\pi P)_j$$

□

On fixe i et on considère $T = T_i = \inf\{n \geq 1 : X_n = i\}$ et

$$\gamma(j) = \gamma^i(j) = \mathbb{E}_i\left[\sum_{0\leq k<T_i} \mathbf{1}_{(X_k=j)}\right]$$

Proposition 4.21. — *On suppose la chaîne irréductible récurrente. Alors*

1. *La mesure γ^i est invariante et de support I, $\forall j$, $\gamma^i(j) > 0$*
2. *Toute mesure invariante est un multiple de γ^i .*

En conséquence, la chaîne admet une probabilité invariante ssi γ^i est de masse finie, et alors l'unique mesure de probabilité invariante est $\pi(j) = \gamma^i(j)/\mathbb{E}_i[T_i]$. En particulier,

$$\pi(i) = \frac{1}{\mathbb{E}_i[T_i]}$$

. On dit que la chaîne est positive récurrente

Démonstration. — 1) Soit $f : I \to \mathbb{R}$ positive. Alors, comme $T_i <= \infty$ et $X_{T_i} = i$, on a

$$\gamma(f) = \sum_j f(j)\gamma(j) = \mathbb{E}_i\left[\sum_{0\le k<T_i} f(X_k)\right] = \mathbb{E}_i\left[\sum_{1\le k<=T_i} f(X_k)\right]$$
$$= \mathbb{E}_i\left[\sum_{1\le k<+\infty} f(X_k)\,\mathbf{1}_{(k\le T_i)}\right]$$
$$= \sum_{1\le k<+\infty} \mathbb{E}_i\left[f(X_k)\,\mathbf{1}_{(k\le T_i)}\right]$$

Or $\{k \le T - i\} \in \mathcal{F}_{k-1}$ donc

$$\mathbb{E}_i\left[f(X_k)\,\mathbf{1}_{(k\le T_i)}\right] = \mathbb{E}_i\left[\mathbb{E}\left[f(X_k) \mid \mathcal{F}_{k-1}\right]\mathbf{1}_{(k\le T_i)}\right] = \mathbb{E}_i\left[Pf(X_{k-1})\,\mathbf{1}_{(k\le T_i)}\right]$$

et donc

$$\gamma(f) \sum_{1\le k<+\infty} \mathbb{E}_i\left[Pf(X_{k-1})\,\mathbf{1}_{(k\le T_i)}\right]$$
$$= \sum_{0\le k<+\infty} \mathbb{E}_i\left[Pf(X_m)\,\mathbf{1}_{(m<T_i)}\right] = \gamma(Pf)$$

ce qui prouve que $\gamma = \gamma P$.
2) En décomposant suivant les temsp de retour succesifs en i, on obtient

$$\mathbb{E}_i\left[\sum_n f(X_n)\right] = \sum_k \gamma^i(f) = +\infty$$

si $\gamma^i(f) > 0$ et 0 sinon. Or pour $f(x) = \mathbf{1}_{(x=j)}$ on sait que $\sum_n f(X_n) = +\infty$ ps, donc $\gamma^i(f) = \gamma^i(j) > 0$.
3) la preuve est pénible et demande des arguments de martingale listés ci dessous □

Remarque. — *Si une chaîne rec irr admet une proba invariante π, alors elle est positive récurrente, pi et γ^i sont proportionnelles, donc γ^i est de masse finie.*

Corollaire 4.22. — *Soit X une chaîne de markov irréductible sur un espace d'états fini. Alors la chaîne est positive récurrente et admet une unique probabilité invariante.*

Lemme 4.23. — *Soit une chaine irréductible récurrente. Si $Pf = f$ et $f \ge 0$, alors f est constante.*

Démonstration. — $f(X_n)$ est une martingale positive, donc converge ps donc ps il existe $c = c(\omega)$ et n_0 tq pour tout $n \geq n_0$, $f(X_n) = c$. Or ps X_n visite tous les points infiniment souvent, donc $\forall i in I, f(i) = c(\omega)$ ce qui prouve que f est constante. $\square$

Lemme 4.24. — *Soit μ une mesure invariante qui charge tout l'espace. Alors $Q(y,x) = \frac{\mu(x)}{\mu(y)} p(x,y)$ définit une matrice stochastique. Soit Y_n une chaîne de Markov de matrice Q. Alors si P est récurrente irréductible, il en est de même de Q. Si π est une mesure invariante pour P et $f = \pi/mu$, alors $Qf = f$.*

Démonstration. — L'irréductibilité de Y est triviale, il suffit de trouver un chemin de y à x dans le graphe de P.
On a évidemment $q^{(n)}(y,x) = \frac{\mu(x)}{\mu(y)} p^{(n)}(x,y)$ et donc $\sum_n q^{(n)}(x,x) = +\infty$ ssi $\sum_n p_{xx}^{(n)} = +\infty$. $\square$

Fin de la preuve de la proposition On a si $f = \pi/\mu$ qui vérife $Qf = f$ donc Q étant irréductible récurrente, f est constante, ie les mesures invariantes sont toutes proportionnelles.

5.1 Théorème ergodique

Théorème 4.25. — *Soit X une chaîne irréductible positive récurrente. Alors, pour toute fonction f positive ou bornée, presque sûrement*

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \pi(f)$$

Démonstration. — C'est le renewal reward theorem avec pour temps de régénération T_i avec i fixé. On a ps

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \gamma^i(f)/\mathbb{E}_i\left[T_i\right] = \pi(f)$$

$\square$

Exemple 4.7. — *Pour l'exemple de l'introduction, la proportion de temps passé en temps long dans l'état 2, en partant de 1, est de 3/8. Elle vaut $\pi(2) = 3/8$ avec π unique proba invariante de la chaîne réduite à $\{1,2,3\}$.*

5.2 Mesures réversibles Il est parfois difficile de calculer la mesure de proba invariante même quand on siat qu'elle existe et est unique.

Définition 4.6. — *On dit que la mesure π est réversible si elle vérifie les équations de bilan détaillé*

$$\pi(x)p_{xy} = \pi(y)p_{yx}$$

Lemme 4.26. — *Une mesure réversible est invariante.*

Démonstration. —
$$\pi P(j) = \sum_i \pi(i)p_{ij} = \sum_i \pi(j)p_{ji} = \pi(j)\sum_i p_{ji} = \pi(j)$$
□

Remarque. — *Si la matrice de transition est symétrique, alors la mesure $\pi(i) = 1$ est réversible donc invariante et si on est sur un espace fini, cela entraîne que l'unique mesure invariante est la loi uniforme.*

Remarque. — *On montre facilement que si$0 \leq N$, et $Y_N = X_{N-n}, 0 \leq n \leq N$ est la chaîne retournée dans le temps (à l'instant N), si $X_0 \sim \pi$ loi réversible, alors Y est une chaîne de Markov (π, P) en montrant que*
$$\mathbb{P}(Y_0 = i_0, \ldots, Y_N = i_N) = \mathbb{P}(X_0 = i_0, \ldots, X_N = i_N)$$

SECTION 6

Convergence vers l'équilibre

La matrice de transition $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ vérifie $P^{2n} = I$ et $P^{2n+1} = P$. En conséquence $p_{ij}^{(n)}$ ne converge pas

Proposition 4.27. — *On considère une chaîne irréductible* apériodique, *i.e. de période $d = 1$. Si I est fini, il existe n_0 tel que pour tout $n \geq n_0$, et tous j, k $p^{(n)}(j,k) > 0$.*

Démonstration. — On pose $S = \left\{n \geq 1, p_{ii}^{(n)} > 0\right\}$ ALors S est un semi groupe de $(\mathbb{N}, +)$ i.e. $S + S \subset S$ et $pgcd(S) = 1$, donc si $I = S - S$, alros I est un idéal de $\mathbb{Z}$ donc de l a forme $d\mathbb{Z}$ et comme $pgcd(S) = 1$ on a $d = 1$ donc $1 \in S - S$: il existe $n_0 \in S$ tel que $n_0 + 1 \in S$ Alors si $k \geq n_0^2 + n_0$, quand on fait la division de k par n_0, $k = qn_0 + r$, on a $q \geq n_0$ donc $k = r(n_0+1) + (q-r)n_0 \in S$. Comme $i \to j$ on en déduit qu'il existe $n(i,j)$ tel que pour tout $n \geq n(i,j)$ $p_{i,j}^{(n)} > 0$. □

Théorème 4.28. — *Soit une chaîne de Markov irréductible positive récurrente de proba invariante π et apériodique. Alors, pour toute loi initiale*
$$\mathbb{P}(X_n = i) \to \pi(i)$$
en particulier pour tout i,
$$p_{ij}^{(n)} \to \pi(j)$$

Démonstration. — Preuve par couplage. Dans le cas d'un espace d'états fini on peut utiliser Perron-Frobenius (voir en appendice). □

Remarque. — *Dans le cas où I est fini, cette convergence s'écrit*

$$P^n \to P^\infty\,, \tag{49}$$

avec P^∞ la matrice dont toutes les lignes sont égales à π. Le théorème de Perron Frobenius assure l'existence d'une constante $C > 0$ telle que

$$\|P^n - P^\infty\| \le C\rho^n \tag{50}$$

avec $\rho = \sup|\lambda| : \lambda \in sp(P), \lambda \neq 1 < 1$: c'est à dire la convergence est exponentiellement rapide.

SECTION 7

Le processus de Galton Watson

On se donne une famille IId $(\xi_i^n, n \geq 1, i \geq 1)$ de variables aléatoires IID de même loi que la variable entière ξ. Etant une variable aléatoire Z_0 entière, indépendante de cette suite, on définit la suite aléatoire $(Z_n)_{n\geq 0}$ par la récurrence:

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{n+1}\,. \tag{51}$$

On dispose donc à chaque génération (instant) n de Z_n individus qui se reproduisent indépendamment suivant la même loi de reproduction, donnant naissance à Z_{n+1} individus à la génération n.
Etant donné que l'on peut écrire

$$Z_{n+1} = F(Z_n, \eta_{n+1}) \tag{52}$$

avec $\eta_n = (\xi_i^n, i \geq 1)$ IID, on a bien une chaîne de Markov homogène sur $\mathbb{N}$.

Je cite Wikipédia
À l'origine, ce modèle a été introduit par Bienaymé en 1845 et indépendamment par Galton en 1873 en vue d'étudier la disparition des patronymes.
Supposons que chaque adulte mâle transmette son patronyme à chacun de ses enfants. Supposons également que le nombre d'enfants de chaque homme soit une variable aléatoire entière (et que la distribution de probabilité soit la même pour tous les hommes dans une lignée). Alors, un patronyme dont les porteurs ont un nombre d'enfant strictement inférieur à 1 en moyenne est amené à disparaître. Inversement, si le nombre moyen d'enfants est supérieur à 1, alors la probabilité de survie de ce nom est non nulle et en cas de survie, le nombre de porteurs du patronyme connaît une croissance exponentielle.

5

Martingales

Les martingales sont les invariants forts des systèmes aléatoires. Elles sont le pendant des lois de conservation des systèmes physiques.

SECTION 1

Filtrations et Martingales

1.1 Définitions

Définition 5.1. — *Un espace probabilisé filtré* $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathbb{P})$ *est la donnée*

- *d'un espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$ *;*
- *d'une* filtration $(\mathcal{F}_n)_n$*, c'est à dire qu'une famille croissante de sous tribus de* $\mathcal{F}$*:*

$$\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \cdots \subset \mathcal{F} .$$

La tribu $\mathcal{F}_n$ représente la quantité d'information connue à l'instant n. Habituellement $\mathcal{F}_n = \mathcal{F}_n^W = \sigma(W_0, \ldots, W_n)$ est la *filtration naturelle* d'un processus W. On note également $\mathcal{F}_\infty = \vee_n \mathcal{F}_n$ la plus petite tribu qui contient toutes les tribus $\mathcal{F}_n$.

Définition 5.2. — *Un processus* $(X_n)_n$ *est* adapté *(à la filtration* $(\mathcal{F}_n)_n$ *si pour tout* n*, la variable aléatoire* X_n *est* $\mathcal{F}_n$*-mesurable. Un processus* $(X_n)_n$ *est dit intégrable si pour tout* n *la variable aléatoire* X_n *est intégrable :* $\forall n$, $\mathbb{E}\left[|X_n|\right] < +\infty$.

Définition 5.3. — *Une* martingale *est un processus adapté intégrable* $(X_n)_n$ *tel que*

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = X_n \qquad (\forall n) .$$

Une sousmartingale *est un processus adapté intégrable* $(X_n)_n$ *tel que*

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n \qquad (\forall n) .$$

Une surmartingale *est un processus adapté intégrable* $(X_n)_n$ *tel que*

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n \qquad (\forall n) .$$

Remarque. — – *Une surmartingale décroît en moyenne.*
- $(X_n)_n$ *est une surmartingale ssi* $(-X_n)_n$ *est une sousmartingale.*
- *Si* $(X_n)_n$ *est une sousmartingale, alors* $(X_n - X_0)_n$ *est une sousmartingale issue de* 0.

1.2 Somme de variables aléatoires indépendantes centrées Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, intégrables, de moyenne nulle : $\mathbb{E}[X_k] = 0$. On pose

$$S_0 = 0, \quad S_n = X_1 + \cdots + X_n\,, \quad \mathcal{F}_n = \sigma(X_1, \ldots, X_n), \quad \mathcal{F}_0 = \{\emptyset, \Omega\}\,.$$

Le processus $(S_n)_n$ est évidemment adapté et intégrable (une somme de variables aléatoires intégrables est intégrable), et

$$\begin{aligned}
\mathbb{E}[S_{n+1} - S_n \mid \mathcal{F}_n] &= \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \\
&= \mathbb{E}[X_{n+1}] \qquad \text{(car } X_{n+1} \text{ est indépendante de } \mathcal{F}_n) \\
&= 0.
\end{aligned}$$

Une question intéressante est de savoir quand la suite S_n converge.
Si $(X_i)_{i\geq 1}$ IID et $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$ c'est la marche aléatoire simple sur $\mathbb{Z}$ et S_n est une martingale.
Si X_i est IID et $\mathbb{P}(X_i = 1) = p = 1 - \mathbb{P}(X_i = -1)$ alors S_n est la marche aléatoire biaisée et c'est $M_n = S_n - np$ qui et une martingale.

1.3 Produit de variables aléatoires indépendantes intégrables de moyenne 1 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, intégrables, de moyenne 1 : $\mathbb{E}[X_k] = 1$. On pose

$$M_0 = 1, \quad M_n = X_1 \cdots X_n\,, \quad \mathcal{F}_n = \sigma(X_1, \ldots, X_n), \quad \mathcal{F}_0 = \{\emptyset, \Omega\}\,.$$

Alors le processus $(M_n)_n$ est évidemment adapté, intégrable (le produit de variables aléatoires indépendantes intégrables est intégrable) et

$$\begin{aligned}
\mathbb{E}[M_{n+1}\mathcal{F}_n] &= \mathbb{E}[M_n X_{n+1}\mathcal{F}_n] \\
&= M_n \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \qquad \text{(car } M_n \text{ est } \mathcal{F}_n \text{ mesurable)} \\
&= M_n \mathbb{E}[X_{n+1}] \qquad \text{(car } X_{n+1} \text{ est indépendante de } \mathcal{F}_n) \\
&= M_n.
\end{aligned}$$

Dans le cas positif, le théorème de convergence des surmartingales entraîne que la martingale positive M_n converge presque sûrement vers une va intégrable M_∞. Nous verrons sous quelles conditions on a $\mathbb{E}[M_\infty] = 1$.
En mathématiques financières, X_n positive représente le changement de valeur (multiplicatif $\pm$ 10 pour cent par exemple) en un période. On prend $X = e^N/\mathbb{E}\left[e^N\right]$ avec $N \sim \mathcal{N}(0,1)$ our un modèle continu, mais le modèle le plus utilisé est le modèle binomial pour lequel $X = (1+a)e^{-v}$ avec la proba p et $X = (1+a)^{-1}e^{-v}$ avec la proba $1-p$ avec donc $e^v = p(1+a) + (1-p)/(1+a)$. Alors M_n est binomiale:

$$\mathbb{P}\left(M_n = X_0(1+a)^m e^{-vn}\right) = \binom{n}{\frac{m+n}{2}} p^{(m+n)/2}(1-p)^{(n-m)/2}$$

Pour la marche aléatoire biaisée $M_n = \rho^{S_n}$ avec $\rho = \frac{1-p}{p}$ est une martingale de type multiplicatif car $\xi_1 = \rho^{X_1}$ vérifie $\mathbb{E}[\xi_1] = p\rho + (1-p)\frac{1}{\rho} = 1$.

1.4 Accumulation de l'information sur une variable aléatoire

Soit $(\mathcal{F}_n)_n$ une filtration et $\xi \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ une variable aléatoire intégrable. On pose $X_n = \mathbb{E}[\xi \mid \mathcal{F}_n]$. C'est évidemment un processus adapté et intégrable. En outre, grâce à la propriété de projection de l'espérance conditionnelle (Proposition 3.3),

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\xi \mid \mathcal{F}_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\xi \mid \mathcal{F}_n] = X_n\,.$$

Dans ce cas, nous serons capable de montrer que $M_n \to M_\infty = \mathbb{E}[\xi \mid \mathcal{F}_\infty]$, c'est à dire que l'information sur ξ connue à l'instant n converge vers l'information maximale sur ξ.
On verra que toute martingale UI est de ce type

1.5 Le processus de GW

Soit $(\xi_i^{(k)}, i \geq 1, k \geq 1)$ iid de même loi que N variable aléatoire entière dont la loi est appelée loi de reproduction : $\xi_i^{(k)}$ est le nombre d'enfants de la personne numéro i de la génération k. Alors le nombre d'individus à l 'instant $n+1$ est

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n+1)}$$

On montre que $W_n = m^{-n}Z^n$ est une martingale avec $m = \mathbb{E}[N] <= \infty$ nombre d'enfants moyen et $\mathcal{F}_n = \sigma(\xi_i^{(k)}, i \geq 1, k \leq n)$.

1.6 Jeu équilibré et stratégie de mise

Définition 5.4. — *Un processus $(C_n)_{n\geq 1}$ est* prévisible *si pour tout $n \geq 1$, la variable aléatoire C_n est $\mathcal{F}_{n-1}$ mesurable. Si $(X_n)_{n\in\mathbb{N}}$ est une sous(sur)martingale,* *alors le processus transformé $(Y_n = (C.X)_n)_{n\in\mathbb{N}}$ est l'intégrale stochastique discrète:*

$$Y_0 = 0\,, \quad Y_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1}) \quad (n \geq 1).$$

La martingale X représente un jeu équilibré, et C la stratégie de mise : on mise C_n à l'instant n, C_n ne dépendant que de l'information à l'instant précédent $n-1$. Alors Y_n représente le gain à l'instant n. Le théorème suivant dit que si le jeu est équilibré, alors il n'y a pas de stratégie de mise gagnante (en moyenne).

Théorème 5.1. — *(a) Soit $(C_n)_{n\geq 1}$ un processus prévisible positif borné (pour tout $n \geq 1$, il existe $K_n < +\infty$ telle que presque sûrement $|C_n(\omega)| \leq K_n$). Soit $(X_n)_{n\in\mathbb{N}}$ une surmartingale. Alors $C.X$ est une surmartingale nulle en zéro.*
(b) Soit $(C_n)_{n\geq 12}$ un processus prévisible borné et X une martingale. Alors, $C.X$ est une martingale nulle en zéro.

Démonstration. — Soit $Y_n = (C.X)_n$. Par construction, comme les variables aléatoires C_k sont bornées et les variables aléatoires X_k intégrables, le processus Y est intégrable ; de même il est évidemment adapté. en outre, comme C_{n+1} est $\mathcal{F}_n$ mesurable,

$$\mathbb{E}\left[Y_{n+1} - Y_n \mid \mathcal{F}_n\right] = \mathbb{E}\left[C_{n+1}(X_{n+1} - X_n) \mid \mathcal{F}_n\right] = C_{n+1}\mathbb{E}\left[(X_{n+1} - X_n) \mid \mathcal{F}_n\right].$$

(a) On a $\mathbb{E}\left[(X_{n+1} - X_n) \mid \mathcal{F}_n\right] \leq 0$ et $C_{n+1} \geq 0$, donc $\mathbb{E}\left[Y_{n+1} - Y_n \mid \mathcal{F}_n\right] \leq 0$.
(b) On a $\mathbb{E}\left[(X_{n+1} - X_n) \mid \mathcal{F}_n\right] = 0$, donc $\mathbb{E}\left[Y_{n+1} - Y_n \mid \mathcal{F}_n\right] = 0$. □

Examinons la stratégie de doublement dans un jeu de pile ou face équilibré. Au début je mise 1, si je perds je mise 2, si je perds encore je mise 4, etc .... On a $C_n = 2^k$ si $X_{n-k} = -1 = ... = X_{n-1}$ j'ai perdu les k fois précédentes, et $C_n = 1$ sinon. On a toujours $\mathbb{E}\left[Y_n\right] = 0$.
Mais on dit qu'en fait on s'arrête de jouer la première fois que l'on gagne et que la le gain est $1 = 2^{k+1} - (1 + 2 + \cdots + 2^k)$. C'est vrai mais la on ne s'arrête pas à un instant fixé d'avance.

SECTION 2

Temps d'arrêt et Théorèmes d'Arrêt

Définition 5.5. — *Un* temps d'arrêt *est une application $T : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ telle que :*

$$\forall n\,, \qquad \{T \leq n\} \in \mathcal{F}_n\,.$$

Remarque. — *T est un temps d'arrêt si pour tout n, $\{T = n\} \in \mathcal{F}_n$. En effet $\{T = 0\} = \{T \le 0\}$, $\{T = n\} = \{T \le n\} \backslash \{T \le n-1\}$ et $\{T \le n\} = \cup_{k=1}^{n} \{T = k\}$.*
Intuitivement, la décision d'arêter à l'instant n est prise uniquement en fonction des informations connues à l'instant n.

Exemple 5.1. — *Étant donnée une suite adaptée $(X_n)_{n\in\mathbb{N}}$, sont des temps d'arrêt:*

$$T = \inf \{n \ge 0 : X_n \in A\} \quad (\inf \emptyset = +\infty) \quad \textit{temps d'entrée dans } A$$
$$T = \inf \{n \ge 1 : X_n \in A\} \qquad \textit{temps de passage dans } A$$
$$T = \inf \{n \ge 1 : |X_n - X_{n-1}| \ge 2\}$$

Définition 5.6. — *Étant donnés un processus $(X_n)_{n\in\mathbb{N}}$ et un temps d'arrêt T, le* processus arrêté en T *X^T est défini par : $X_n^T = X_{T\wedge n}$.*

Si $(X_n)_{n\in\mathbb{N}}$ est adapté, alors pour tout n, $X_{T\wedge n}$ est $\mathcal{F}_{T\wedge n}$ mesurable, et donc $\mathcal{F}_n$ mesurable : le processus arrêté est adapté.

Proposition 5.2. — *Une sousmartingale arrêtée est une sousmartingale.*

Démonstration. — Soit donc $(X_n)_{n\in\mathbb{N}}$ une sousmartingale, T un temps d'arrêt et le processus $C_n = \mathbf{1}_{(T\ge n)}$. Alors,

$$(C.X)_n = \sum_{k=1}^{n} \mathbf{1}_{(T\le k)}(X_k - X_{k-1}) = X_{T\wedge n} - X_0$$

c'est à dire $C.X = X^T - X_0$. Le processus $(C_n)_{n\ge 1}$ est positif borné et prévisible car $\{T \ge n\}^C = \{T \le n-1\} \in \mathcal{F}_{n-1}$. Donc $C.X$, et par conséquence X^T, est une sousmartingale. □

Théorème 5.3 **(*Théorème de convergence des sousmartingales*)**
Soit $(X_n)_{n\in\mathbb{N}}$ une sousmartingale telle que $\sup_n \mathbb{E}\left[X_n^+\right] < +\infty$. Alors, $(X_n)_{n\in\mathbb{N}}$ est bornée dans L^1, i.e. $\sup_n \mathbb{E}\left[|X_n|\right] < +\infty$, et X_n converge presque sûrement vers une variable aléatoire X_∞ qui est intégrable.

Corollaire 5.4. — *Une surmatingale positive converge presque sûrement vers une variable aléatoire intégrable.*

Corollaire 5.5. — *Soit $(X_n)_{n\in\mathbb{N}}$ une sousmartingale bornée dans L^p avec $p > 1$. Alors $\sup_n |X_n|$ est dans L^p et X_n converge presque sûrement et dans L^p vers une variable aléatoire X_∞.*

Remarque. — *Ce résultat est faux pour $p = 1$. On peut montrer, à l'aide du théorème de Kakutani sur les martingales produits, que si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires indépendantes positive de même loi non dégénérée,* *telle que* $\exp X_1 = 1$*, alors* $M_n = X_1 \dots X_n$ *est une martingale, bornée dans* L^1 *car* $\mathbb{E}[M_n] = 1$*, et* M_n *converge presque sûrement vers* $M_\infty = 0$*.*

Théorème 5.6 **(*Théorème de convergence des martingales*)**
Soit $(X_n)_{n\in\mathbb{N}}$ *une martingale. Sont équivalentes,*

(i) $(X_n)_{n\in\mathbb{N}}$ *est Uniformément intégrable.*
(ii) $(X_n)_{n\in\mathbb{N}}$ *converge dans* L^1*.*
(iii) Il existe une variable aléatoire intégrable X_∞*,* $\mathcal{F}_\infty$ *mesurable, telle que* $X_n = \mathbb{E}[X_\infty \mid \mathcal{F}_n]$*.*

Dans ces conditions X_n *converge aussi vers* X_∞ *presque sûrement et dans* L^1*. On dit que la martingale est fermée, ou régulière.*

On obtient des énoncés semblables pour les sur et sous martingales UI.
Le *Théorème d'arrêt* suivant signifie que pour une martingale Uniformément intégrable, on peut remplacer les temps fixes par des temps d'arrêt quelconques, même infinis.

Théorème 5.7*. — Soit* $(X_n)_{n\in\mathbb{N}}$ *une martingale Uniformément intégrable. Pour tout temps d'arrêt* T *on définit:* $M_T = M_n$ *si* $T(\omega) = n$*,* $M_T = M_\infty$*, si* $T(\omega) = +\infty$*. Alors,*

(i) $X_T = \mathbb{E}[X_\infty \mid \mathcal{F}_T]$
(ii) $\mathbb{E}[X_0] = \mathbb{E}[X_T]$*.*

Kolmogorov a établi des conditions nécessaires et suffisantes pour la convergence d'une série de variables aléatoires indépendantes $\sum_n X_n$.

Théorème 5.8 **(*Théorème des trois séries*)***. — Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires indépendantes. Alors* $\sum_n X_n$ *converge presque sûrement si et seulement si il existe une constante* $K > 0$ *(et alors ce sera vrai pour toutes les constantes* $K > 0$*), telle que*

(i) $\sum_n \mathbb{P}(|X_n| > K) < +\infty$
(ii) $\sum_n \mathbb{E}\left[X_n^K\right]$ *converge*
(iii) $\sum_n \mathrm{Var}(X_n^K) < +\infty$

avec la notation : $X^K(\omega) = X(\omega)\,\mathbf{1}_{(|X(\omega)|\leq K)}$*.*

6

Appendices

SECTION 1

Le lemme de Doob

Théorème 6.1. — *Soit $f : E \to (F, \mathcal{F})$. On note $\sigma(f) = f^{-1}(\mathcal{F})$ la plus petite tribu sur E qui rend f mesurable. Soit $g : (E, \sigma(f)) \to (G, \mathcal{G})$ avec $\mathcal{G}$ la tribu des boréliens de G, avec G espace métrique séparable. Alors il existe une fonction $h : (F, \mathcal{F}) \to (G, \mathcal{G})$ mesurable telle que $g = h \circ f$.*

Démonstration. — On traite le cas $G = [0, 1]$ puis on utilise le fait qu'un espace métrique séparable est mesurablement isomorphe à une partie de $[0, 1]$ muni de sa tribu des boréliens.
Si $g = 1_A$ est une indicatrice, alors $A \in \sigma(f)$ donc $A = f^{-1}(B)$ avec $B \in \mathcal{F}$ et $g = 1_B \circ f$.
Par linéarité, si g est une fonction étagée $g = \sum_i \alpha_i 1_{A_i}$ avec $A_i = f^{-1}(B_i)$ alors, $g = h \circ f$ avec $h = \sum_i \alpha_i 1_{B_i}$.
Si g est $\sigma(f)$ mesurable positive, alors il exite une suite de fonctions croissantes étagées positives g_n convergeant vers g. Il exites h_n mesurable positive telle que $g_n = h_n \circ f$. L'ensemble D des x pour lesquels la suite h_n converge vers une limite finie est mesurable, et on pose $h(x) = (\limsup_n h_n(x))\, \mathbf{1}_{(x \in D)}$. C'est bien une fonction mesurable positive et par construction $f(x) \in D$, car $h_n(f(x)) \to g(x)$, donc $g = h \circ f$. □

SECTION 2

Le théorème des classes monotones

2.1 Classes monotones et Tribus

2.1.1 Classes monotones Ω étant un ensemble, un ensemble de parties $\mathcal{M}$ de Ω est appelée *classe monotone* si les propriétés suivantes sont vérifiées
i) $\Omega \in \mathcal{M}$
ii) si $A, B \in \mathcal{M}$ et si $B \subseteq A$ alors $A \backslash B \in \mathcal{M}$
iii) si $A_j \in \mathcal{M}, \forall j \in \mathbb{N}$ et si $A_j \subseteq A_{j+1}$ alors $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{M}$.

Proposition 6.2. — *Si $\mathcal{E}$ est un ensemble de parties de Ω, l'intersection des classes monotones contenant $\mathcal{E}$ est une classe monotone, notée $\mathcal{M}(\mathcal{E})$, appelée classe monotone engendrée par $\mathcal{E}$.*

L'intérêt de cette notion est dans le résultat suivant

Théorème 6.3. — *Si $\mathcal{E}$ est une famille de parties de Ω stable par intersection finie, alors la classe monotone engendrée par $\mathcal{E}$ coïncide avec la tribu engendrée par $\mathcal{E}$.*

Démonstration. — Soit $\sigma(\mathcal{E})$ la tribu engendrée par $\mathcal{E}$. Une tribu étant une classe monotone on a l'inclusion $\mathcal{M}(\mathcal{E}) \subseteq \sigma(\mathcal{E})$.
Suivant l'Exercice.1 il suffit alors de montrer que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie.
Définissons

$$\mathcal{M}_1 = \{A \in \mathcal{M}(\mathcal{E}),\ \forall B \in \mathcal{E},\ A \cap B \in \mathcal{M}(\mathcal{E})\}$$

$\mathcal{M}_1$ est une classe monotone contenant $\mathcal{E}$ donc $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}_1$.
Ensuite définissons

$$\mathcal{M}_2 = \{A \in \mathcal{M}(\mathcal{E}),\ \forall B \in \mathcal{M}(\mathcal{E}),\ A \cap B \in \mathcal{M}(\mathcal{E})\}$$

On vérifie de même que $\mathcal{M}_2$ est une classe monotone. Pour conclure il suffit de montrer qu'elle contient $\mathcal{E}$. Soient $A \in \mathcal{E}$ et $B \in \mathcal{M}(\mathcal{E})$. Or on sait que $B \in \mathcal{M}_1$, on a donc $B \cap C \in \mathcal{M}(\mathcal{E})$ pour tout $C \in \mathcal{E}$. Il suffit de choisir $C = A$. □

Le théorème des classes monotones admet une version fonctionnelle qu'il est utile de connaitre et qui s'énonce comme suit.

Définition 6.1. — *Un ensemble $\mathcal{F}$ de fonctions réelles sur Ω est dit stable par convergence monotone bornée si pour toute suite bornée et monotone $\{f_n\}$ déléments de $\mathcal{F}$ la limite simple sur Ω, $f(x) = \lim_{n \to +\infty} f_n(x)$ est dans $\mathcal{F}$.*
Un espace vectoriel monotone est un espace vectoriel de fonctions rélles et bornées sur Ω, contenant les constantes et stable par convergence monotone bornée

Il est facile de voir que si μ, ν sont deux probabilités, l'ensemble des fonctions mesurables bornées f telles que $\int f d\mu = \int f d\nu$ est un e.v. monotone.

Théorème 6.4 **(classes monotones fonctionnelles)**
Soit $\mathcal{F}$ un espace vectoriel monotone. Si $\mathcal{C}$ est une partie de $\mathcal{F}$ stable par multiplication et contenant les constantes alors $\mathcal{F}$ contient toutes les fonctions bornées et $\sigma(\mathcal{C})$-mesurables.

Démonstration. Introduisons $\mathcal{F}_0$ le plus petit sous-espace vectoriel de $\mathcal{F}$ contenant $\mathcal{C}$ et stable par convergence monotone bornée.
Montrons que $\mathcal{F}_0$ est stable par multiplication. Soient $f \in \mathcal{C}$ et $g \in \mathcal{F}_0$. Montrons alors que $fg \in \mathcal{F}_0$. Il suffit de supposer que $f \geq 0$. L'argument qui suit est semblable à celui utilisé dans 6.3.
Introduisons :

$$\mathcal{F}_1 = \{g \in \mathcal{F}_0,\ fg \in \mathcal{F}_0\}.$$

$\mathcal{F}_1$ est un espace vectoriel, contenant $\mathcal{C}$ (et les constantes) et stable par convergence monotone bornée. Donc $\mathcal{F}_0 \subseteq \mathcal{F}_1$ et on a bien montré que $fg \in \mathcal{F}_0$.
Montrons maintenant que si $g, h \in \mathcal{F}_0$ alors $gh \in \mathcal{F}_0$.
Pour cela on introduit

$$\mathcal{F}_2 = \{h \in \mathcal{F}_0,\ fg \in \mathcal{F}_0\}.$$

$\mathcal{F}_2$ est un espace vectoriel contenant les constantes et quitte à faire une translation on peut supposer que $g \geq 0$. $\mathcal{F}_2$ est également stable par convergence monotone bornée et contient $\mathcal{C}$. On a donc $\mathcal{F}_0 \subseteq \mathcal{F}_2$.
Montrons maintenant que $\mathcal{F}_0$ coïncide avec l'ensemble des fonctions bornées et $\sigma(\mathcal{C})$-mesurable sur Ω.
Soit $f \in \mathcal{F}_0$. Montrons que $|f| \in \mathcal{F}_0$. Il en résultera que si $f, g \in \mathcal{F}_0$ alors $\min(f, g) \in \mathcal{F}_0$ et $\max(f, g) \in \mathcal{F}_0$.
On obtient cette propriété en utilisant une suite croissante de polynomes $p_n(x)$sur $[-1, 1]$ telle $\lim\limits_{n\to+\infty} p_n(x) = |x|$. (Exercice)
On pose $\mathcal{A} = \{A \subseteq \Omega,\ 1_A \in \mathcal{F}_0\}$. On vérifie facilement grâce aux propriétés précédentes que $\mathcal{A}$ est une tribu. Notons par $\mathcal{B}(\mathcal{A})$ l'espace des fonctions réelles, bornées et $\mathcal{A}$-mesurables. Soit $f \in \mathcal{B}(\mathcal{A}$. Si de plus $f \geq 0$, on sait (voir appendice, (M6)) qu'il existe une suite croissante de fonctions $\mathcal{A}$-étagées telle que $\lim\limits_{n\to+\infty} f_n = f$ simplement sur Ω. On en déduit donc que $f \in \mathcal{F}_0$. D' où il en résulte que $\mathcal{B}(\mathcal{A} \subseteq \mathcal{F}_0$. Montrons que l'on a l'égalité. Soit $f \in \mathcal{F}_0$. $\mathcal{F}_0$ étant un espace vectoriel contenant les constantes, il suffit de supposer que $f \geq 0$ et de montrer que $[f \geq 1] \in \mathcal{A}$. Or la suite $g_n = (\min\{1, f\})^n$ est une suite monotone de $\mathcal{F}_0$ qui converge vers $1_{[f\geq 1]} \in \mathcal{F}_0$.
On en déduit que $\sigma(\mathcal{C}) \subseteq \mathcal{A}$ (car si $f \in \mathcal{C}$ alors $f \in \mathcal{F}_0$ est donc $\mathcal{A}$-mesurable).
Par suite on a

$$\mathcal{B}(\sigma(\mathcal{C})) \subseteq \mathcal{B}(\mathcal{A}) = \mathcal{F}_0 \subseteq \mathcal{F}.$$

Donnons une application du Théorème précédent.

Corollaire 6.5. — *Soient E un espace métrique muni de sa tribu borélienne et μ, μ' deux mesures finies sur E. Alors $\mu = \mu'$ si et seulement si pour toute fonction réelle f continue et bornée sur E on a $\int f d\mu = \int f d\mu'$.*

Démonstration. — On utilise pour $\mathcal{C}$ l'algèbre $\mathcal{C}_b(E)$ des fonctions continues et bornées sur E. Il est facile de montrer que la tribu engendrée par $\mathcal{C}_b(E)$ coïcide avec la tribu borélienne de E. Ce qui donne le résultat. □

Remarque. — *Lorsque $E = \mathbb{R}^n$ on peut également utiliser $\mathcal{C} = \mathcal{S}(\mathbb{R}^n$, $\mathcal{S}(\mathbb{R}^n$ étant l'espace de Schwartz (voir appendice).*

SECTION 3

Perron Frobenius Theorem

This theorem gives sufficient conditions for matrices with positive entries to have their spectral radius as the only eignevalue for a vector with strictly positive entries.
For $x \in \mathbb{R}^d$ we say that $x \geq 0$ if forall i, $x_i \geq 0$. We say $x > 0$ if $x \geq 0$ and $x \neq 0$, and $x >> 0$ if forall i, $x_i > 0$. We have the same notations for a real matrix $A \in \mathcal{M}_{d\times d}$. Observe that if $x > 0$ and $A >> 0$ then $Ax >> 0$.

Lemme 6.6. — *If $x >> 0$ and $y > 0$ then $a = \sup\{c > 0 : x \geq cy\} > 0$ (with the convention $\sup \emptyset = -\infty$) and $\exists i$ s.t. $x_i = ay_i$.*

Démonstration. — Let $c = \frac{\inf x_i}{\sup y_j} > 0$. Then, $x \geq cy$, so $a > 0$. We have $z = x - ay \geq 0$. If $z >> 0$, then there exists $\delta > 0$ s.t. $z \geq \delta y$ and so $x \geq (a + \delta)y$ and this contradicts the definition of a. So there exists i such that $z_i = 0$. □

Lemme 6.7. — *If $A >> 0$, $x > 0$ and $Ax \geq cx$ then $c \leq \rho(A)$ the spectral radius.*

Démonstration. — Assume $c > 0$. We have by induction $A^n x \geq c^n x$, therefore

$$\|A^n\| = \sup\left\{\frac{\|A^n y\|_\infty}{\|y\|_\infty}, y \neq 0\right\} \geq \frac{\|A^n x\|_\infty}{\|x\|_\infty} \geq c^n \,, \tag{53}$$

and by Gelfand's formula, $\rho(A) = \lim_{n\to+\infty} \|A^n\|^{1/n} \geq c$. □

Théorème 6.8. — *Assume $A >> 0$ and let $\rho = \rho(A) > 0$ be its spectral radius. Then*

1. *There exists $u >> 0$ such that $Au = \rho u$.*
2. *If $\lambda \in sp(A)$ and $\lambda \neq \rho$ then, $|\lambda| < \rho$.*
3. *The dimension of the eigenspace associated to ρ is 1. More precisely, if $Ax = \lambda x$ with $x > 0$ and $\lambda > 0$, then $\lambda = \rho$ and $x = \alpha u$ for some $\alpha > 0$.*

Démonstration. — Let $x \neq 0$ be an eigenvector of A with eigenvalue λ. Then $y = |x|$ defined by $y_i = |x_i|$ satisfies

$$|\lambda| y_i = \left| \sum_j a_{ij} x_j \right| \leq \sum_j a_{ij} y_j \tag{54}$$

that is $Ay \geq |\lambda| y$. By Lemma 6.7, $|\lambda| \leq \rho$.
Assume that $|\lambda| = \rho$. Then by the two preceding lemmas, $\rho = \sup \{c > 0 : Ay \geq cy\}$ and there exits i such that $\rho y_i = (Ay)_i$ that is there is equality in (**??**)

$$\rho y_i = \left| \sum_j a_{ij} x_j \right| = \sum_j a_{ij} y_j = \sum_j a_{ij} |x_j| \, . \tag{55}$$

Recall that if $w_1, \ldots, w_n$ are complex numbers such that $|w_1 + \cdots + w_n| = \sum |w_i|$ then they have the same argument : there exists $\theta \in \mathbb{R}$ and $\lambda_i \geq 0$ such that $w_k = \lambda_k e^{i\theta}$.
Applying this to the preceding equation (55) yields that the x_j have the same argument, and therefore $x_j = e^{i\theta} y_j$ so $x = e^{i\theta} y$ and thus $Ay = \lambda y$. We also get that $y_i > 0$ since otherwise forall j $y_j = 0$. Since $\rho y_i = (Ay)_i$ we get that $\lambda = \rho$. Since $y > 0$ and $A >> 0$ we have $y = \frac{1}{\lambda} Ay >> 0$ and this proves statements 1 and 2 of the theorem.
Let us prove now that if $Ax = \lambda x$ with $x > 0$ and $\lambda > 0$, then $\lambda = \rho$ and $x = au$ for some $a > 0$. First $x > 0$, $A >> 0$ so $x = \frac{1}{\lambda} Ax >> 0$.
Let $a = \sup \{c > 0 : x \geq cu\}$. By Lemma 6.6, $a > 0$ and for some i, $x_i = au_i$. Therefore

$$\lambda a u_i = \lambda x_i = (Ax)_i = \sum_j a_{ij} x_j \geq_{(x \geq au)} a \sum_j a_{ij} u_j = a \rho u_i$$

SInce $|\lambda| \leq \rho$, we obtain $\lambda = \rho$ and that there is equality in the preceding inequality, and therefore since $a_{ij} > 0$, $x = \alpha u$. □

Théorème 6.9. — *If $A \geq 0$ and there exists $m \in \mathbb{N}^*$ such that $A^m >> 0$ then the conclusions of the preceding theorem hold.*

Démonstration. — We have $\rho(A^m) = \rho(A)^m$. Let $n \neq 0$ and λ such that $Ax = \lambda x$. Then $A^m x = \lambda^m x$ If $|\lambda| = \rho$ then $|\lambda^m| = \rho(A^m)$ so by the preceding theorem $\lambda^m = \rho$.
Observe that since $A^m >> 0$, $0 \notin sp(A)$ so every line and every column of A is > 0. Therefore $A^{m+1} = A^m A >> 0$ and we have also $\lambda^{m+1} = \rho$. This yields $\lambda = \rho$ so $A^m x = \rho^m x$ and by the preceding theorem $x = \eta y$ with $y >> 0$, $\eta \in \mathbb{C}$. Therefore $Ay = y$ and we are done. □

Assume $A \geq 0$ and $A^m >> 0$ for some integer m. We saw that A has a right eigenvector $u >> 0$ such that $Au = \rho u$. We normalize u so that $\sum_i u_i = 1$.

Then A has a left eigenvector with eigenvalue ρ (consider the transpose): $vA = \rho v$. And we can normalize v so that $vu = u.v^T = \sum_i v_i u_= 1$.
Prove as an exercice that the operator $Px = (x.v^T)u = vxu$ is a projector $P^2 = P$ that commutes with A : $AP = PA = \rho P$. (It is a projector on the eigenspace with eigenvalue ρ).

Lemme 6.10. — *Let $B = A - \rho P$. Then $\rho(B) < \rho$ and $\frac{A^n}{\rho^n} \to P$.*

Démonstration. — Assume $Bx = \lambda x$ with $\lambda \neq 0$, $x \neq 0$. Then

$$\lambda Px = PBx = PAx - \rho P^2 x = 0\,, \tag{56}$$

o $Px = 0$ and $Ax = \lambda x$. If $|\lambda| = \rho$ then $\lambda = \rho$ and $Px = x$ so $x = 0$, absurd. Therefore $|\lambda| < \rho$ and we have proved that $\rho(B) < \rho$.By induction

$$A^n = B^n + \rho^n P\,, \tag{57}$$

and therefore $\rho^{-n} A^n = P + \rho^{-n} B^n$. But if $\delta > 0$, then there exists by Gelfand formula a n_0 such that for $n \geq n_0$

$$\left\| \rho^{-n} B^n \right\| \leq \rho{-}n(\rho B + \delta)^n \to 0 \tag{58}$$

if we have chosen δ small enough. And thus $\rho^{-n} B^n \to 0$.

□

Prove as an exercice that if $x > 0$ and $Ax \geq \lambda x$ for some $\lambda \geq 0$ then $\lambda = \rho$ and so x is a multiple of u.

Fortunately, we know exactly when to apply Theorem 6.9. A matrix $A \geq 0$ is *irreducible* if

$$\forall x, y \quad \exists m \quad (A^m)_{xy} > 0\,. \tag{59}$$

The *period* of an element x is $d(x) = gcd\{n \geq 1 : (A^n)_{xx} > 0\}$. If A is irreducible then all states have the same period, forall x $d(x) = d(A)$. We say then that A is *aperiodic* if $d(A) = 1$.

Proposition 6.11. — *Let A be a matrix such that $A \geq 0$. Then there exists $m \in \mathbb{N}^*$ such that $A^m >> 0$ iff A is irreducible and aperiodic.*

7

Exercices

◇ Exercice 1.1

Soit N_1, N_2 deux variables aléatoires indépendantes qui suivent des lois de Poisson de paramètres respectifs λ_1 et λ_2, et soit $N = N_1 + N_2$.

1. Déterminer la loi de N.
2. Montrer que conditionnellement à $N = k$, avec $k \in \mathbb{N}^*$, la loi de N_1 est une binomiale $\mathcal{B}(n, p)$ dont on déterminera les paramètres n et p.

◇ Exercice 1.2

Des clients arrivent dans un magasin au rythme de 3 a la minute en moyenne. Les temps ‘ d'arrivée suivent un processus de Poisson.

1. Calculer la probabilité qu'il arrive un seul client en 2 minutes.
2. Calculer la probabilité qu'il n'arrive aucun client en 5 minutes.
3. Calculer la probabilité pour que, dans 2 périodes de 2 minutes, disjointes, il arrive au moins 2 clients.

◇ Exercice 1.3

Un compteur de particules n'enregistre qu'une particule sur deux arrivant au compteur. On suppose que les particules arrivent selon un processus de Poisson, au rythme moyen de 4 par minute. Soit S le temps d'attente entre deux particules enregistrées.

1. Calculer la loi de S, $\mathbb{E}[S]$ et $\mathrm{Var}(S)$.

2. Calculer la probabilité $\mathbb{P}(S \geq 1)$.

⋄ Exercice 1.4

Soient N un processus de Poisson d'intensité $\lambda > 0$ et X_0 une variable aléatoire indépendante de N de loi $\frac{1}{2}(\delta_1 + \delta_{-1})$. On pose $X_t = X_0(-1)^{N_t}$, pour tout $t \geq 0$. Calculer :

1. $\mathbb{E}[X_t]$ pour tout $t \geq 0$.
2. $\mathrm{Cov}(X_s, X_t)$ pour $0 < s \leq t$.

⋄ Exercice 1.5

Soit N un processus de Poisson d'intensité $\lambda > 0$. On pose pour tout entier $n \geq 1$, $U_n = \frac{N_n}{n}$ et $V_n = \sqrt{\frac{n}{\lambda}}(U_n - \lambda)$.

1. Calculer les transformées de Laplace de U_n et V_n.
2. En déduire les limites en loi de U_n et V_n lorsque n tend vers l'infini.

⋄ Exercice 1.6

Soit $(X_t)_{t\geq 0}$ et $(Y_t)_{t\geq 0}$ deux processus de Poisson indépendantes, de paramètres respectifs λ et μ. Montrer que le nombre d'arrivées du processus Y se produisant entre deux arrivées successives du processus X suit la loi de $T-1$ avec T une variable géométrique de paramètre $\frac{\lambda}{\lambda+\mu}$.

⋄ Exercice 1.7

Les arrivées de clients dans un magasin forment un processus de Poisson de taux 20 par heure. a) Calculez la probabilité que trois clients, exactement, arrivent pendant la première heure d'ouverture du magasin. b) Déterminez la distribution du nombre de ventes sur une journée de 8 heures, sachant que la probabilité qu'un client achète quelque chose est égale à 0.3. Justifiez soigneusement vos réponses.

⋄ Exercice 1.8

Remplacement minimal Une machine a une durée de vie sans pane D qui est une variable aléatoire admettant pour densité une fonctio continue f. Elle est remplacée régulièrement aux instants $T, 2T, 3T, \ldots.$ En outre une réparation minimale est effectuée pour chaque panne se produisant entre deux remplacements programmés. Cette réparation prend un temps négligeable.

1. Décrire le processus qui compte le nombre total de réparations au cours du temps, en justifiant au moins heuristiquement vos hypothèses.
2. Calculer la loi du nombre total X de réparations entre deux remplacements. Déeminer $m = \mathbb{E}[X]$.
3. Application numérique. On suppose $T = 10$. Calculer m dans les deux cas suivants.
   - D suit une loi exponentielle standard.
   - D suit une loi uniforme sur $[0, 2T]$.

⋄ Exercice 1.9

Boum la girafe Boum la girafe désire traverser la route. A l'endroit où elle désire traverser, les automobiles passent suivant un processus de Poisson de paramètre $\lambda = 1/3mn^{-1}$. Boum met $c = 10$ secondes pour traverser.

1. Calculer la probabilité que Boum se fasse percuter (et porte bien son nom).
2. Gaston le Héron met lui aussi c secondes pour traverser, mais il a une bonne vue, et attend de voir un intervalle de temps suffisant avant de traverser, c'est à dire qu'il attend $J = \inf\{i \geq 1, X_i > c\}$ pour traverser avec X_i les intervalles de passage et $S_n = X_1 + \cdots + X_n$. Montrer que le temps moyen de passage est $\mathbb{E}[S_{J-1} + c] = (e^{\lambda c} - 1)/\lambda$.

⋄ Exercice 1.10

Soit $(\tau_i, i \geq 1)$ une famille de variables aléatoires indépendantes, avec $\tau_i \sim \mathcal{E}(\beta i)$, avec $\beta > 0$ donné. On note $T_n = \tau_1 + \cdots + \tau_n$ et on désire montrer que pour tout $t > 0$, $\mathbb{P}(T_n \leq t) = (1 - e^{\beta t})^n$.

1. Soit $M_n = \max(Y_1, \dots, Y_n)$ avec $(Y_i, i \geq 1)$ IID de loi $\mathcal{E}(\beta)$. Déterminer la loi de M_n.
2. On considère n processus de Poisson, mutuellement indépendants, de même paramètre β et $Y_1, \cdots, Y_n$ leurs premiers temps de saut ainsi que leur statistiques d'ordre $0 < Y^{(1)} < Y^{(2)} < \cdots < Y^{(n)}$. On note σ_i les intervalles $\sigma_1 = Y^{(1)}$ et $\sigma_i = Y^{(i)} - Y^{(i-1)}$ pour $i \geq 2$. Montrer que $(\sigma_1, \dots, \sigma_n) \stackrel{d}{=} (\tau_n, \dots, \tau_1)$.
3. En déduire que $T_n \stackrel{d}{=} M_n$.

$\diamond$ Exercice 1.11

Un pêcheur attrape les truites en suivant un processus de Poisson de 3 truites par heure. La truite pèse en moyenne 2 kilos avec un écart type de 0,5 kilos. Donner la moyenne et la variance du poids de truites pêchées en 2 heures.

$\diamond$ Exercice 1.12

Les clients se présentent devant un distributeur automatique de billets en suivant un Processus de Poisson non homogène i.e. de taux $\lambda(t)$ dépendant du temps. Entre 8h et 10h, λ est constante et vaut 5. Entre 10 heures et 14 heures, le taux s'accroît linéairement de 5 par heure pour atteindre 25 par heure à 14 heures. Entre 14 heures et 20 heures, le taux décroît linéairement pour atteindre 4 par heures à 20 heures. Entre 20 heures et 24 heures, le taux vaut 3.
Les montants retirés par les clients sont des variables aléatoires indépendantes de moyenne 100 euros et d'écart type 125 euros.

1. Quelle est la loi du nombre de clients qui retirent de l'argent entre 8 heures et 24 heures ?
2. Déterminer l'espérance est la variance du montant total X retiré pendant cette période.
3. Majorer du mieux possible $\mathbb{P}(X \geq 25000)$.

$\diamond$ Exercice 2.1

On considère un buffer(tampon) de capacité infinie. Les messages arrivent suivant un processus de Poisson de paramètre $\lambda > 0$. On suppose le tampon de capacité infinie, et que l'on décide de vider le tampon toutes les T unités de temps. Le coût du vidage de tampon est de K et le coût de stockage d'un message dans le tampon est de h par unité de temps. Quelle est la valeur de T pour laquelle le coût moyen en temps long est le plus petit ? Un cycle a pour durée déterministe T. Le temps d'occupation pour la totalité des messages dans un cycle est un processus de Poisson composé

$$\tau_T = \sum_{i=1}^{N(T)} (T - S_i)$$

1. Utiliser la formule exponentielle pour les processus de Poisson pour montrer que
$$\mathbb{E}\left[e^{-\alpha\tau_T}\right] = \exp\left(-\lambda T(1 - \frac{1 - e^{-\alpha T}}{\alpha T})\right).$$

2. Montrer que $\mathbb{E}[\tau_T] = \frac{1}{2}\lambda T^2$ et donc le coût moyen par cycle $g(\lambda, T, K)$ (pour $h = 1$).
3. En déduire la valeur de la taille de tampon optimale $T* = \sqrt{\frac{2K}{\lambda h}}$.

$\diamond$ Exercice 2.2

La lampe d'un réverbère est remplacée lorsqu'elle est défaillante, et également à des instants programmés :T,2T, 3T, ... Il y a toujours remplacement, quel que soit l'âge réel de la lampe. Les durées de vie des lampes sont des variables aléatoires de loi $\gamma(2, \mu)$. Quel est le nombre moyen de lampes utilisées dans une période $]kT, (k+1)T]$?

$\diamond$ Exercice 2.3

Un service de limousine entre la gare et l'aéroport fonctionne du matin jusqu'au soir. Les limousines quittent la gare à des intervalles de temps dont la longueur est uniformément distribuée entre 10 et 20 minutes. Vous êtes supposé arriver à la gare à 3 heures de l'après-midi. Essayez d'estimer la moyenne et la variance du temps d'attente qu'une limousine parte pour l'aéroport ?

$\diamond$ Exercice 2.4

Les commandes parviennent à un atelier suivant un processus de renouvellement de moyenne $\frac{1}{\lambda}$. La production ne commence que lorsque N commandes sont en attente. Le temps de production est négligeable. Un coût fixe de $K > 0$ est nécessaire pour lancer un cycle de production et si une commande est mise en attente j unité de temps, cela coûte hj.

1. Identifier les instants de renouvellement.
2. Déterminer le coût moyen par unité de temps, en temps long.
3. Quelle valeur de N minimise ce coût ?
4. Supposons maintenant que l'on change la politique de production en attendant un temps fixe $T > 0$ pour lancer la production. Comment cela change-t-il la valeur du coût moyen par unité de temps, en temps long ? Calculer ce coût moyen si les commandes arrivent suivant un processus de Poisson de paramètre λ et optimisez T.

$\diamond$ Exercice 2.5

Un bateau mouche part toutes les T minutes, $T > 0$ fixé. Les clients potentiels arrivent suivant un processus de Poisson de paramètre $\lambda > 0$. Si un client potentiel voit que le prochain bateau par dans s minutes, il achète un billet avec la probabilité $e^{-\mu s}, 0 \leq s \leq T$ avec $\mu > 0$ paramètre fixé. On suppose que le bateau a une capacité infinie.

1. Quelle est la nature du processus des clients (qui achètent un billet) ?
2. Quelle est la loi du nombre de passager par bateau ?
3. Supposons que le coût d'une promenade est K euros et que le prix du billet est R euros. Identifier les instants de renouvellement et calculer le gain moyen par unité de temps. Montrer que l'optimum à K, R, μ, λ fixés est obtenu pour une valeur t^*.
4. Que peut on dire maintenant si le temps qu'un client potentiel est prêt à attendre ne suit plus une loi exponentielle de paramètre μ ?

$\diamond$ Exercice 3.1

Soit X une variable aléatoire qui s'écrit $X = f(Y, Z)$ où les variables aléatoires Y et Z sont indépendantes. Montrez que la loi conditionnelle de X sachant $Y = y$ est la loi de la variable aléatoire $f(y, Z)$.
On suppose maintenant que $X = f(Y, Z)$ où la variable aléatoire est $\mathcal{G}$ mesurable, et où la variable aléatoire Z est indépendante de $\mathcal{G}$. Déterminez la loi conditionnelle de X sachant $\mathcal{G}$.

$\diamond$ Exercice 3.2

Soit (X, Y) deux variables aléatoires réelles telles que $(X, Y) \stackrel{d}{=} (Y, X)$ (on dit qu'elles sont échangeables en loi). On suppose X intégrable. Montrer qu'il existe une fonction mesurable f telle que

$$\mathbb{E}\left[X \mid Y\right] = f(Y) , \quad \text{et} \quad \mathbb{E}\left[Y \mid X\right] = f(X) .$$

$\diamond$ Exercice 3.3

Soit $X_1, \cdots, X_n$ des variables aléatoires indépendantes de même loi $\mathcal{N}(0, 1)$ et soit $S_n = X_1 + \cdots X_n$. Montrez que la loi de X_1 sachant $S_n = s$ est $\mathcal{N}(s/n, 1 - \frac{1}{n})$.

$\diamond$ Exercice 3.4

Soit $X_1, \ldots, X_p$ des variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs $\lambda_1, \ldots, \lambda_p$. Déterminer la loi conditionnelle du vecteur aléatoire $(X_1, \ldots, X_p)$ sachant $\sum_i X_i = n$. (On peut commencer par traiter le cas $p = 2$).

$\diamond$ Exercice 3.5

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$. Etant donnée une variable aléatoire de carré intégrable X on pose $\mathrm{Var}(X \mid \mathcal{G}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid \mathcal{G}])^2 \mid \mathcal{G}\right]$. Montrer que

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X \mid \mathcal{G})] + \mathrm{Var}(\mathbb{E}[X \mid \mathcal{G}])$$

$\diamond$ Exercice 3.6

Soit X, Y deux variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(1,p)$ i.e. $\mathbb{P}(X = 1) = 1 - \mathbb{P}(X = 0) = p$ avec $0 < p < 1$. On pose $Z = \mathbf{1}_{(X+Y=0)}$ et $\mathcal{G} = \sigma(Z)$. Calculer $\mathbb{E}[X \mid \mathcal{G}]$ et $\mathbb{E}[Y \mid \mathcal{G}]$. Ces deux variables aléatoires sont-elles indépendantes en général ?

$\diamond$ Exercice 3.7

Soit (X, Y) un couple de variables aléatoires admettant pour densité

$$f(x,y) = 4y(x-y)e^{-(x+y)}\,\mathbf{1}_{(0 \le y \le x)}$$

1. Déterminer la loi conditionnelle de X sachant Y.
2. En déduire $\mathbb{E}[X \mid Y]$
3. Calculer $\mathbb{P}(X < 1 \mid Y)$.

$\diamond$ Exercice 3.8

Soit X et Y deux variables aléatoires rélles. La loi de X admet pour densité

$$\frac{2}{(\log 2)^2} \frac{\log(1+x)}{x}\,\mathbf{1}_{(0<x<1)}$$

La loi conditionnelle de Y sachant $X = x$, avec $0 < x < 1$ admet pour densité

$$\mathbf{1}_{(0<y<x)} \frac{1}{(1+y)\log(1+x)}$$

1. Les variables aléatoires X et Y sont elles indépendantes ?
2. Quelle est la loi de Y ?

3. Quelle est la loi de X sachant $Y = y$?
4. Déterminer $\mathbb{E}[Y \mid X]$

◇ Exercice 3.9

Soit X une variable aléatoire réelle de loi $\mathcal{E}(\lambda)$ avec $\lambda > 1$ et soit $Y = \lfloor X \rfloor$ sa partie entière.

1. Déterminer la loi conditionnelle de $X - Y$ sachant Y.
2. En déduire $\mathbb{E}[X \mid Y]$

◇ Exercice 4.1

Pour le climat à Nantes $P = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$. Montrer que si $X_0 = 1$ ps, ie on part d'un jour où il fait beau, alors

$$\mathbb{P}(X_n = 1) = \frac{1}{2}(1 + 2^{-n})$$

(ne pas oublier que $\mathbb{P}(X_n = 2) = 1 - \mathbb{P}(X_n = 1)$). Qu'advient il de la loi de X_n quand n est grand ?

◇ Exercice 4.2

On considère une chaîne de Markov homogène de matrice de transition

$$P = \begin{pmatrix} 7/20 & 3/20 & 1/4 & 1/4 \\ 3/10 & 1/4 & 7/20 & 1/10 \\ 1/4 & 1/4 & 7/20 & 3/20 \\ 3/10 & 1/4 & 1/4 & 1/5 \end{pmatrix}$$

1. Quelle est la loi conditionelle de X_{62} sachant $X_{60} = 3$.
2. On suppose que la chaîne est issue de la loi uniforme $\lambda = (1/4, 1/4, 1/4, 1/4)$. Déterminer la loi à l'instant 2.

◇ Exercice 4.3

Deux villes A et B totalisent une population de 4300 habitants. La ville A est plus agréable, mais la ville B offre de meilleures perspectives de carrière. On observe que :

- 78 % des habitants de A habitent encore en A l'année suivante,
- 8 % des habitants de B partent chaque année habiter A.

On modélise l'évolution du lieu d'habitation d'un individu au cours des années par une chaîne de Markov homogène $(X_n)_{n\in\mathbb{N}}$.

1. Ecrire la matrice de transition.
2. Calculer la probabilité que Dupont réside dans la ville A l'année 2008 sachant qu'il était dans la ville B en 2006.
3. Si ce modèle de mouvement entre les villes A et B s'applique depuis longtemps à toutes les personnes des deux villes, c'est à dire à un million de personnes, le nombre d'habitants de la ville B devrait être de N, nombre à déterminer.

◇ Exercice 4.4

Identifier les classes de communicationde la matrice de transition

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

Quelles classes sont fermées ? Quels sont les états récurrents et les états transients ? Quelles sont les probabilités invariantes ?

◇ Exercice 4.5

Un jeu de l'oie simple est constitué de 10 cases numérotées de 1 à 10 disposées comme ci-dessous. La case 1 est la case de départ et la case 10 est la case d'arrivée.

12345678910

Pour faire avancer le pion, on lance un dé à 6 faces numérotées de 1à 6 et on avance le pion d'un nombre de cases égal au nombre obtenu avec le dé. Le jeu s'arrête lorsque le pion tombe exactement sur la case 10. Sinon le pion recule. Par exemple, si le pion se trouve sur la case 9 et si le dé tombe sur 3, le pion va à la case 8. Si au coup suivant, le dé tombe sur 1, le pion retourne sur la case 9.
Les positions successives du pion définissent une cha ne de Markov sur les entiers de 1 à 10. On supposera que lorsque le jeu s'arr te, les positions suivantes du pion sont toujours 10.
Donner la matrice de transition de cette chaîne de Markov.

$\diamond$ Exercice 4.6

On considère un marché des télecoms à 3 opérateurs : A,B, C. A l'annéee $n = 0$, les clients sont répartis de la façon suivante: A : 12%, B 40% et C 48%. Chaque année un client peut changer d'opérateur. On modélise son attitude par une chaîne de Markov homogène de matrice de transition

$$P = \begin{pmatrix} 0.1 & 0.3 & 0.6 \\ 0.1 & 0.5 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$$

1. Quelle est sur le long terme la répartition des clients entre les différents opérateurs ?
2. On suppose qu'un contrat souscrit auprès de l'opérateur A est de 500 euros, et qu'un contrat souscrit auprès de B ou C est de 1000 euros. Le marché étant composé de 10 millions de personnes, calculer le profit moyen annuel escompté, au bout de 10 ans, par chaque opérateur.

$\diamond$ Exercice 4.7

On considère une chaîne de Markov homogène dont les états sont numérotés de 1 à 3. Sa matrice de transition est :

$$\begin{pmatrix} 2/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 \end{pmatrix}$$

Monter que lorsque $n \to +\infty$ la quantité $\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{(X_k \geq 2)}$ converge en probabilité vers une constante. Quelle est la valeur de cette constante ?

$\diamond$ Exercice 4.8

JC est en prison et a 3 euros. Il peut sortir en versant une caution de 8 euros. Un garde accepte de faire une série de paris avec lui. Si JC parie A euros, il gagne A euros avec la probabilité $\alpha = 0.4$ et perd A euros avec la probabilité $1 - \alpha$.

1. Trouvez la probabilité qu'il gagne 8 dollars avent de perdre tout son argent si
   (a) Il mise 1 euro à chaque fois
   (b) Il mise à chaque fois le maximum, mais pas plus que le nécessaire pour obtenir 8 euros.
2. Quelle stratégie donne à JC la plus grande chance de sortir de prison ?

3. Ecrire un programme qui simule la chaîne de Markov, et comparez vos résultats théoriques aux simulations (en utilisant la loi des grands nombres bien sûr).

$\diamond$ Exercice 4.9

On considère une chaîne de Markov homogène dont les états sont numérotés de 1 à 5. Sa matrice de transition est :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 3/20 & 1/10 & 3/10 & 1/5 & 1/4 \\ 1/5 & 3/20 & 3/20 & 7/20 & 3/20 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Déterminer les classes de communication de cette chaîne.
2. Calculer la probabilité que la chaîne de Markov, issue de l'état 4, réussisse à atteindre l'état 2.
3. Calculer cette même probabilité en supposant cette fois que la loi initiale de la chaîne est
$$\lambda = (0, 0, 1/3, 1/3, 1/3, 0)\,.$$

$\diamond$ Exercice 4.10

Chaque matin un étudiant prend un des trois livres qu'il possède sur son étagère. La probabilité qu'il choisisse le livre numéro i est $0 < \alpha_i < 1$ et les choix des jours successifs sont indépendants. Le soir, il replace le livre au bout à gauche de l'étagère.

1. Modéliser le processus d'évolution de l'ordre des livres comme une chaîne de Markov X_n à valeurs dans $\mathcal{S}_3$ les permutations de $\{1, 2, 3\}$.
2. Montrer qu'il exite une unique probabilité invariante π.
3. Montrer que $\forall n \geq 1$, $\mathbb{P}\left(X_n(1) = i\right) = \alpha_i$
4. Montrer que si $n \geq 2$, et p_n est la probabilité que le jour n l'étudiant trouve les livres dans l'ordre $1, 2, 3$, alors $p_n = \alpha_1 p_{n-1} + \alpha_1\alpha_2$.
5. En déduire que $p_n \to p$ un nombre que l'on calculera en fonction des α_i.

$\diamond$ Exercice 4.11

Un robot collecteur de boîtes de soda peut se trouver dans 3 états suivant son niveau de batterie:

– 1 : batterie chargée

– 2 : batterie faible
– 3 : batterie à plat

Lorsque la batterie est chargée, elle se trouve faible la période suivante avec la probabilité α et chargée avec la probabilité $1 - \alpha$. Lorsque la batterie est faible, elle se trouve à plat la période suivante avec la probabilité 4α et faible avec la probabilité $1 - 4\alpha$ (on suppose $0 < \alpha < 1/4$). Lorsque la batterie est à plat, on recharge le robot qui se retrouve ensuite avec sa batterie chargée.

1. Décrire l'évolution du robot par une chaîne de Markov dont on écrira explicitement la matrice de transition P. Quelle sont les classes de communication ? Montrer qu'il existe une unique probabilité invariante π que l'on déterminera en fonction de α.
2. On suppose que le coût de recharge de la batterie est $c = -3$ euros, et que par unité de temps, lorsqu'il n'est pas à plat, le robot collecte des boîtes pour une valeur de r euros. Montrer que $g_n = \frac{1}{n}S_n = \frac{1}{n}(R(X_0) + \cdots + R(X_n))$, gain temporel moyen sur n périodes, converge quel que soit l'état de départ vers une quantité $gm(r, \alpha)$.
3. Montrer qu'il existe une valeur $r_c(\alpha)$ telle que $gm(r, \alpha) > 0$ si et seulement si $r > r_c(\alpha)$. C'est la valeur de collecte à partir de laquelle il est rentable d'installer le robot.

⋄Exercice 4.12

On considère la chaîne de Markov $(X_n)_{n\geq 0}$ sur $I = \{1, 2, 3, 4\}$ de matrice de transition

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

1. Montrer que la chaîne est irréductible, positive récurrente.
2. Trouver toutes les probabilités invariantes.
3. Montrer que $\frac{1}{n}(X_0 + X_1 + \ldots + X_{n-1})$ et $\frac{1}{n}(X_0^2 + X_1^2 + \ldots + X_{n-1}^2)$ convergent presque sûrement et déterminer leurs limites.

⋄Exercice 4.13

Une chanteuse d'opéra doit donner une longue série de concerts. Son tempérament d'artiste la pousse à vouloir tous les soirs arrêter les concerts, et ce avec une probabilité de $\frac{1}{2}$. Une fois qu'elle a décidé d'arrêter, elle ne chantera pas de nouveau jusqu'à ce que l'organisateur la convainque de son admiration. Pour cela il lui envoie des fleurs chaque jour jusqu'à ce qu'elle revienne. Des fleurs coutant x milliers d'euros, $0 \leq x \leq 1$, amènent une réconciliation avec la probabilité $\sqrt{x}$. L'organisateur fait 750 euros de bénéfice à chaque représentation donnée. Combien doit-il dépenser en fleurs ?

$\diamond$ Exercice 4.14

Un parieur possède 2 euros et désire obtenir 10 euros le plus vite possible. Il peut jouer au jeu suivant. Une pièce équilibrée est lancée. S'il parie sur le bon côté, il gagne une somme égale à sa mise, sinon il perd sa mise. Le parieur décide d'une statégie assez simpliste dans laquelle il mise tout son argent s'il a moins de 5 euros, et juste assez pour obtenir 10 euros s'il gagne sinon.
Soit $X_0 = 2$ et X_n son capital après n lancers. Montrez que le parieur obtiendra ses 10 euros avec une probabilité de $1/5$. Quel est le nombre moyen de lancers avant que le parieur ne perde tout ou parte avec son capital de 10 euros ? Pouvez vous trouver une meilleure stratégie ?

$\diamond$ Exercice 4.15

On considère une chaîne de Markov sur l'espace d'états a, b, c, d de matrice de transition

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

1. Calculer la probabilité que la chaîne soit dans l'état b après 10 transitions, sachant qu'après 8 transitions elle est dans l'état a.
2. Montrer que la chaîne admet une unique probabilité invariante π que l'on déterminera.
3. Calculer pour tout état initial i, le temps moyen $k_i = \mathbb{E}_i [H_a]$ d'atteinte de l'état a : $H_a = \inf \{n \geq 0 : X_n = a\}$.
4. Calculer le temps moyen de retour en a, $\mathbb{E}_a [T_a]$ avec $T_a = \inf \{n \geq 1 : X_n = a\}$. Vérifier votre calcul en utilisant les résultats de la question 3.
5. On suppose que on vous propose un jeu pour lequel lorsque l'on se trouve dans l'état i, on se voit payer la somme $f(i)$ euros, avec $f(a) = 0, f(b) = f(c) = 1, f(d) = -2$. Devez vous accepter de jouer ?

$\diamond$ Exercice 4.16

On entraîne une souris à choisir un objet A parmi deux objets A et B en répétant l'expérience suivante : on lui présente les deux objets; si elle choisit

A on lui donne de la nourriture, sinon elle n'a rien. A chaque étape elle peut être dans un des trois états d'esprit suivant : dans l'état 1,elle ne peut pas se rappeller quel objet procure la récompense,et elle choisit au hasard ; dans l'état 2, elle se rappelle, choisit A, mais elle peut de nouveau oublier ; dans l'état 3 elle se rappelle, choisit A, et n'oublie jamais. Après chaque expérience, elle peut changer d'état d'esprit suivant la matrice de transition

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{12} & \frac{5}{12} \\ 0 & 0 & 1 \end{pmatrix}$$

Elle est dans l'état 1 avant la première expérience. Quelle est la probabilité qu'elle soit dans l'état 1 juste avant la $(n+1)$-ième expérience ? Quelle est la probabilité $p_{n+1}(A)$ qu'elle choisisse A à la $(n+1)$-ième expérience ? Est ce que la suite $p_{n+1}(A)$ converge ?

$\diamond$ Exercice 4.17

On considère une machine outil qui perce des trous dans une pièce de métal. Elle met 45 secondes pour percer un trou. Lorsqu'elle est bien réglée, elle perce toujours le trou parfaitement au centre de la pièce. Malheureusement, après chaque opération, elle a deux pour cent de chances de se dérégler. Lorsqu'elle est déréglée, elle perce un trou sur cinq hors du centre, et elle reste déréglée. Heureusement, on scanne toutes les pièces, et si on a un trou décentré, on rerègle la machine (on la réaligne), une opération qui prend 45 secondes : pendant cette période la machine ne perce aucun trou bien sûr.

1. Modéliser la vie de la machine outil par une chaîne de Markov sur un espace à 3 états à choisir parmi les quatre états suivants:
    - a: la machine est bien réglée et elle vient de percer un trou centré
    - b: la machine est bien réglée et elle vient de percer un trou décentré
    - c: la machine est déréglée et elle vient de percer un trou centré
    - d: la machine est déréglée et elle vient de percer un trou décentré
2. On suppose que l'on fait tourner cette machine sur un temps très long. Quelle est la proportion de pièces défectueuses produites ? Quel est le nombre moyen de pièces défectueuses produites sur une période d'une heure ?
3. Répondre aux questions précédentes, en supposant que le temps de réglage est maintenant de six minutes.

$\diamond$ Exercice 4.18

On considère une chaîne de Markov homogène dont les états sont numérotés de 1 à 4. Sa matrice de transition est :

$$\begin{pmatrix} 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \\ * & * & 0 & 0 \end{pmatrix}$$

où $*$ désigne un nombre > 0. Donner la période de l'état numéro 2.

⋄ Exercice 4.19

On considère un processus de Galton-Watson $(Z_n)_{n\in\mathbb{N}}$ de loi de reproduction N.

1. Montrer que si $Z_0 = x$ et si $G(t) = \mathbb{E}\left[t^N\right]$ est la fonction génératrice de N, alors la fonction génératrice de Z_n est $G_{Z_n}(t) = (G^{(n)}(t))^x$, avec $G^{(n)} = G \circ \cdots \circ G$ la composée n ième de la fonction G.
2. Montrer que si Z et Z' sont deux processus de Galton-Watson indépendants de même loi de reproduction N, issus de x et y, alors $Z''_n = Z_n + Z'_n$ est un processus de Galton-Watson de loi de reproduction N, issu de $x + y$.(Ceci est appelée propriété de branchement).
3. En déduire que $\mathbb{P}_i(Z_n = 0) = \mathbb{P}_1(Z_n = 0)^i = q_n^i$ et donc que la probabilité d'extinction $q = \lim q_n$ vérifie
$$\mathbb{P}_i(\exists n, Z_n = 0) = q^i \,.$$
4. Ecrire le système d'équations satisfaites par $h_i = \mathbb{P}_i(\exists n, Z_n = 0)$ et en déduire que q est solution de $G(q) = q$. Cela donne une autre démonstration du résultat vu en cours : il y a extinction presque sûre ssi $\mathbb{E}[N] \leq 1$ et $\mathbb{P}(N = 1) < 1$.

⋄ Exercice 5.1

Soit $(X_n)_{n\in\mathbb{N}}$ une surmartingale positive et soit

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

avec la convention $\inf \emptyset = +\infty$.

1. Montrez que presque sûrement : $\forall n \geq T\,,\ X_n = 0$.
2. Montrez que presque sûrement $X_\infty = \lim X_n$ existe et que l'on a presque sûrement
$$\{X_\infty > 0\} \subset \{\forall n, X_n > 0\,,\}\,,$$

et que l'on a égalité si les variables aléatoires X_n sont à valeur dans $\mathbb{N}$. Pouvez vous trouver un exemple où il y a inclusion stricte ?

$\diamond$ Exercice 5.2

(Urne de Polya). Une urne contient N boules noires et B boules blanches. Une boule est tirée au hasard, suivant une probabilité uniforme sur les boules dans l'urne. Elle est remise dans l'urne, et on ajoute aussi a boules de la couleur tirée. On itère cette procédure de tirage-ajout. Soit $X_0 = \frac{N}{N+B}$ la proportion d boules noires initialement dans l'urne, et soit X_k la proportion de boules noires à la k ième étape du procédé.

1. Montrer que X_k est une martingale pour sa filtration naturelle $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$.
2. Montrer que cette martingale converge presque sûrement vers une variable aléatoire U, la proportion aléatoire à l'équilibre, et que pour tout entier $p \geq 1$:
$$\mathbb{E}\left[X_k^p\right] \to \mathbb{E}\left[U^p\right].$$
3. On suppose que $N = B = 1$ et $a = 1$. Fixons un entier $k \geq 1$, et considérons $Y_n = (n+2)X_n$ le nombres de boules noires à l'étape n et
$$Z_n = \frac{Y_n(Y_n+1)\dots(Y_n+k-1)}{(n+2)(n+3)\dots(n+k+1)}.$$
Montrez que Z_n est une martingale qui converge presque sûrement. En déduire la valeur de $\mathbb{E}\left[U^k\right]$, puis la loi de la variable U.
4. Pouvez vous traiter le cas général ?

$\diamond$ Exercice 5.3

Soit X_n la marche aléatoire symétrique simple sur $\mathbb{Z}$: $X_0 = 0$, et $X_n = \xi_1 + \cdots \xi_n$ avec ξ_n indépendantes de même loi $\mathbb{P}(\xi_i = \pm 1) = \frac{1}{2}$. On note, pour $a \in \mathbb{Z}$,
$$T_a = \inf\{n \geq 1 : X_n = a\}$$
avec la convention $\inf \emptyset = +\infty$. Soit a un entier,$a \geq 1$, $T = T_a$ $S = T_a \wedge T_{-a}$. On pose $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$

1. Montrer que X_n et $M_n = X_n^2 - n$ sont des martingales. En déduire que $\mathbb{E}[S] = a^2$.
2. Montrer que $Z_n = e^{\lambda X_n - n \log \operatorname{ch}(\lambda)}$ est une martingale.
3. On suppose $\lambda > 0$. Montrer que $Z_{n\wedge T}$ converge presque sûrement et dans L^2 vers la variable aléatoire $W = \frac{e^{\lambda a}}{(\operatorname{ch}(\lambda))^T} \mathbf{1}_{(T<+\infty)}$.

4. En déduire que $T < +\infty$ presque sûrement, puis montrer que
$$\mathbb{E}\left[(\mathrm{ch}(\lambda))^{-T}\right] = e^{\lambda a}\,.$$
5. Montrer que $\mathbb{E}[T] = +\infty$.
6. Montrer que $U_n = \frac{\mathrm{ch}(\lambda X_n)}{(\mathrm{ch}(\lambda))^n}$ est une martingale. En déduire que
$$\mathbb{E}\left[(\mathrm{ch}(\lambda))^{-S}\right] = \frac{1}{\mathrm{ch}(\lambda a)}\,.$$
Retrouver la relation $\mathbb{E}[S] = a^2$.

⋄ Exercice 5.4

Soit Y_n une suite de variables aléatoires définies par

- $Y_0 = 1$;
- sachant $Y_n = y$, Y_{n+1} est uniformément distribuée sur l'intervalle $[0, y]$.

1. Montrer que $X_n = 2^n Y_n$ est une martingale
2. En déduire que presque sûrement la suite $2^n Y_n$ converge vers une limite positive. Peut-on assurer que cette limite est strictement positive ?

⋄ Exercice 5.5

On rappelle que si X_n est une martingale et que pour un $p > 1$, $\sup_n \mathbb{E}[|X_n|^p] < +\infty$, alors X_n converge presque sûrement et dans L^p vers une variable X_∞.

On considère une suite de variables aléatoires définie par $X_0 = \mu \in]0,1[$ et la récurrence
$$X_{n+1} = \begin{cases} \alpha + \beta X_n & \text{avec la probabilité } X_n; \\ \beta X_n & \text{avec la proba } 1 - X_n. \end{cases}$$
avec $\alpha, \beta > 0$ et $\alpha + \beta = 1$.

1. Montrer que X_n est une martingale telle que $X_n \in]0,1[$.
2. Montrer que $\mathbb{E}[X_n] = \mu$ et $\mathrm{Var}(X_n) = (1 - (1-\alpha^2)^n)\mu(1-\mu)$.
3. En déduire que X_n converge presque sûrement vers une variable aléatoire X_∞ telle que $\mathbb{E}[X_\infty] = \mu$ et $\mathrm{Var}\, X_\infty = \mu(1-\mu)$.
4. Montrer que X_∞ suit la loi de Bernoulli $\mathcal{B}(1,\mu)$:
$$\mathbb{P}(X_\infty = 1) = 1 - \mathbb{P}(X_\infty = 0) = \mu$$
.

$\diamond$ Exercice 5.6

On suppose que X_n est une martingale et que $(C_n)_{n\geq 1}$ est un processus prévisible tels que pour un $p \geq 1, \frac{1}{p} + \frac{1}{q} = 1$ on ait

$$\forall n, \quad X_n \in L^p \qquad \forall n \geq 1, \quad C_n \in L^q$$

Montrer que le processus transformé $C.X$ est une martingale nulle en zéro.

$\diamond$ Exercice 5.7

Soit U une variable aléatoire de loi uniforme sur l'intervalle $[0,1]$ et soit $X_0 = 1$, $X_n = n\,\mathbf{1}_{(U\leq 1/n)}$. Montrer que X_n est une martingale pour sa filtration naturelle, que X_n est bornée dans L^1 et que presque sûrement X_n converge vers 0.

$\diamond$ Exercice 5.8

Soit T une variable aléatoire entière telle que pour un $a > 0$, et tout $n \in \mathbb{N}$: $\mathbb{P}(T = n) = a(1+a)^{-n-1}$.

1. On note $\mathcal{F}_n = \sigma(T \wedge n)$ la tribu engendrée par la variable aléatoire $T \wedge n$. Démontrer que $\mathcal{F}_n$ est engendrée par une partition finie, puis que $\mathcal{F}_n$ est une filtration.
2. Etablir les relations
$$\mathbb{E}\left[\,\mathbf{1}_{(T\geq n+1)} \mid \mathcal{F}_n\right] = (1+a)^{-1}\,\mathbf{1}_{(T\geq n)}$$
$$\mathbb{E}\left[T \wedge (n+1) \mid \mathcal{F}_n\right] = T \wedge n + (1+a)^{-1}\,\mathbf{1}_{(T\geq n)}\,.$$
3. Déterminer un réel λ tel que la suite $X_n = \lambda(T \wedge n) + (1+a)^{-1}\,\mathbf{1}_{(T\geq n)}$ soit une martingale. Déterminer alors la décomposition de Doob de X_n^2.

$\diamond$ Exercice 5.9

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi normale $\mathcal{N}(0,\sigma^2)$.On pose $\mathcal{F}_n = \sigma(X_1,\ldots,X_n)$, $\mathcal{F}_0 = \{\emptyset,\Omega\}$, $S_n = X_1 + \cdots + X_n$.

1. Montrer que pour tout u réel, $Z_n^u = \exp(uS_n - \frac{1}{2}nu^2\sigma^2)$ est une martingale.
2. Montrer que pour tout u réel, Z_n^u converge presque sûrement vers une variable aléatoire Z_∞ finie.
3. Pour quelles valeurs de u est ce que Z_n^u est une martingale uniformément intégrable ?

UNIVERSITÉ DE NANTES

E.C.N.

INTRODUCTION AUX PROCESSUS STOCHASTIQUES — ANNÉE 2015/2016

EXAMEN DE JANVIER 2016

LES SEULS DOCUMENT ADMIS SONT LES NOTES DE COURS. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE ET ORDINATEUR NON ADMIS.

Exercice 1

On admet que la population des individus infecté par le VIH croît selon un processus de Poisson d'intensité inconnue $\lambda > 0$. On notera N_t le nombre d'individus infectés à l'instant t, on ne tiendra pas compte des décès.
Chaque individu infecté subit une période d'incubation entre le moment où il est infecté par le virus et le moment où les symptômes du SIDA apparaissent. La durée de cette période d'incubation est aléatoire. Les périodes d'incubation pour les différents individus sont des variables aléatoires IID de loi commune ν de support $\mathbb{R}^+$ et de fonction de répatition G.
A l'instant $t > 0$ on note N_t^1 le nombre d'individus qui présentent les symptômes du SIDA, et N_t^2 le nombre d'individus infectés par le virus, mais ne présentent pas encore les symptômes. On a bien sûr : $N_t = N_t^1 + N_t^2$.
Montrer que N_t^1 et N_t^2 sont deux variables aléatoires indépendantes de loi de Poisson de paramètres respectifs $\lambda\int_0^t G(s)\,ds$ et $\lambda\int_0^t (1-G(s))\,ds$.

Exercice 2

On considère une chaîne de Markov sur l'espace d'états a, b, c, d de matrice de transition

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

1. Calculer la probabilité que la chaîne soit dans l'état b après 10 transitions, sachant qu'après 8 transitions elle est dans l'état a.
2. Montrer que la chaîne admet une unique probabilité invariante π que l'on déterminera.
3. Calculer pour tout état initial i, le temps moyen $k_i = \mathbb{E}_i[H_a]$ d'atteinte de l'état a : $H_a = \inf\{n \geq 0 : X_n = a\}$.
4. Calculer le temps moyen de retour en a, $\mathbb{E}_a[T_a]$ avec $T_a = \inf\{n \geq 1 : X_n = a\}$. Vérifier votre calcul en utilisant les résultats de la question 3.
5. On suppose que on vous propose un jeu pour lequel lorsque l'on se trouve dans l'état i, on se voit payer la somme $f(i)$ euros, avec $f(a) = 0, f(b) = f(c) = 1, f(d) = -2$. Devez vous accepter de jouer ?

Exercice 3

Pour modéliser l'évolution d'une configuration génétique, on considère la chaîne de Markov d'espace d'états $E = \{0, 1, \ldots, N\}$ et de matrice de transition

$$P(i,j) = C_N^j \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}.$$

En d'autres termes, à i fixé, $P(i,\cdot)$ est une loi binômiale $\mathscr{B}(N,i/N)$ (N est la taille de la population et i le nombre d'individus portant un certain caractère génétique).

1. Est-ce-que la chaîne est irréductible ? Quels sont les états récurrents ?
2. Montrer que X_n est une martingale pour la filtration $\mathscr{F}_n = \sigma(X_0,\ldots,X_n)$.
3. Montrer que la limite $\lim_n X_n = X_\infty$ existe ps.
4. Déterminez la loi de X_∞.
5. On pose $Y_n = \frac{X_n}{N}$. Montrer qu'il existe $\rho \in]0,1[$ tel que $Z_n = \rho^{-n} Y_n(1-Y_n)$ soit une martingale.
6. En déduire que X_n converge vers X_∞ au moins exponentiellement vite.

UNIVERSITÉ DE NANTES

E.C.N.

INTRODUCTION AUX PROCESSUS STOCHASTIQUES — ANNÉE 2015/2016

EXAMEN DE MARS 2016

LES SEULS DOCUMENT ADMIS SONT LES NOTES DE COURS. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE ET ORDINATEUR NON ADMIS.

Exercice 1

Soit $(Y_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi, à valeur dans $\mathbb{Z}$, intégrables, telles que $m = \mathbb{E}[Y_1] < 0$, $\mathbb{P}(Y_1 = 1) > 0$ et $\mathbb{P}(Y_1 \geq 2) = 0$. On pose $X_0 = 0$, $X_n = Y_1 + \cdots + X_n$ et $W = \sup_{n\geq 1} X_n$. Le but de ce problème est de déterminer la loi de W.

1. Montrer que $W < +\infty$ presque sûrement (utiliser la loi des grands nombres).
2. Soit X une variable aléatoire réelle et $\psi(\lambda) = \log \mathbb{E}\left[e^{\lambda X}\right]$, avec la convention $\psi(\lambda) = +\infty$ lorsque la transformée de laplace est égale à $+\infty$. Montrer que ψ est convexe (utiliser l'inégalité de Hölder).
3. On pose ici $\psi(\lambda) = \log \mathbb{E}\left[e^{\lambda Y_1}\right]$. Montrer que $\psi(\lambda) < +\infty$ pour tout $\lambda > 0$. Que vaut $\psi'(0+)$ la dérivèe à droite de ψ en 0 ? Montrer que $\psi(\lambda) \to +\infty$ quand $\lambda \to +\infty$ et en déduire l'existence d'un unique $\lambda_0 > 0$ tel que $\psi(\lambda_0) = 0$.
4. Soit $\lambda_0 > 0$ tel que $\psi(\lambda_0) = 0$. Montrer que $Z_n = e^{\lambda_0 X_n}$ est une martingale et établir que $Z_n \to 0$ presque sûrement.
5. Pour k entier, $k \geq 1$ on note $\tau_k = \inf\{n \geq 1 : X_n \geq k\}$. Montrer que
$$\lim Z_{n\wedge\tau_k} = e^{\lambda_0 k}\, \mathbf{1}_{(\tau_k < +\infty)}$$
6. Calculer $\mathbb{P}(\tau_k < +\infty)$ et en déduire la loi de W. Déterminer précisément cette loi dans le cas où $\mathbb{P}(Y_i = 1) = 1 - \mathbb{P}(Y_i = -1) = p < \frac{1}{2}$.

Exercice 2

On suppose que les instants de passage $(T_i, i \in \mathbb{Z})$ des véhicules sur un autoroute en un point a sont répartis suivant une mesure de Poisson sur $\mathbb{R}$ d'intensité μ. On suppose que les vitesses des véhicules $(V_i, i \in \mathbb{Z})$ sont IID de même loi à support dans $]0, +\infty[$, i.e. $\mathbb{P}(V_1 > 0) = 1$. On se donne un point d'observation $b \neq a$. On suppose que les vitesses sont constantes au cours du trajet et que les véhicules peuvent toujours se doubler sans changer de vitesse, ce qui fait que les instants d'observations des mêmes véhicules au point b sont les $T'_i = T_i + \frac{b-a}{V_i}$.

1. Montrer que $(T'_i, i \in \mathbb{Z})$ est un processus de Poisson sur $\mathbb{R}$ d'intensité ν à préciser en fonction de μ et de la loi ρ de V_1.
2. Montrer que si $\mu(dt) = \lambda dt$ est un multiple de la mesure de Lebesgue sur $\mathbb{R}$, alors $\nu = \mu$ (et donc ne dépend pas de la loi des vitesses).
3. Montrer que à l'instant 0 les véhicules sont répartis sur la voie suivant un processus de Poisson d'intensité η à préciser en fonction de μ et de la loi de V_1.
4. Faire le calcul de η dans le cas $\mu = \lambda dt$

UNIVERSITÉ DE NANTES

E.C.N.

INTRODUCTION AUX PROCESSUS STOCHASTIQUES — ANNÉE 2016/2017

EXAMEN DE JANVIER 2017

LES SEULS DOCUMENT ADMIS SONT LES NOTES DE COURS MANUSCRITES. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE ET ORDINATEUR NON ADMIS.

Exercice 1

On suppose que les instants de passage $(T_i, i \in \mathbb{Z})$ des véhicules sur un autoroute en un point a sont répartis suivant une mesure de Poisson sur $\mathbb{R}$ d'intensité μ. On suppose que les vitesses des véhicules $(V_i, i \in \mathbb{Z})$ sont IID de même loi à support dans $]0, +\infty[$, i.e. $\mathbb{P}(V_1 > 0) = 1$. On se donne un point d'observation $b \neq a$. On suppose que les vitesses sont constantes au cours du trajet et que les véhicules peuvent toujours se doubler sans changer de vitesse, ce qui fait que les instants d'observations des mêmes véhicules au point b sont les $T'_i = T_i + \frac{b-a}{V_i}$.

1. Montrer que $(T'_i, i \in \mathbb{Z})$ est un processus de Poisson sur $\mathbb{R}$ d'intensité ν à préciser en fonction de μ et de la loi ρ de V_1.
2. Montrer que si $\mu(dt) = \lambda dt$ est un multiple de la mesure de Lebesgue sur $\mathbb{R}$, alors $\nu = \mu$ (et donc ne dépend pas de la loi des vitesses).
3. Montrer que à l'instant 0 les véhicules sont répartis sur la voie suivant un processus de Poisson d'intensité η à préciser en fonction de μ et de la loi de V_1.
4. Faire le calcul de η dans le cas $\mu = \lambda dt$

Exercice 2

On considère la matrice de transition sur l'espace d'états $S = \{A, B, C, D\}$

$$P = \begin{pmatrix} 0.3 & 0.5 & 0.1 & 0.1 \\ 0.4 & 0.1 & 0.3 & 0.2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Déterminer les classes de communication de la chaîne et leur nature.
2. On suppose la chaîne issue de A. Déterminer le temps moyen pour se trouver pour la première fois dans un des états $\{B, C, D\}$.
3. On suppose la chaîne issue de B. Calculer le temps moyen passé avant d'être absorbée ?
4. Calculer la probabilité que la chaîne issue de A soit absorbée en D ?
5. Montrer que P^k converge vers une matrice P^∞ que l'on calculera explicitement.

UNIVERSITÉ DE NANTES

E.C.N.

INTRODUCTION AUX PROCESSUS STOCHASTIQUES — ANNÉE 2016/2017

EXAMEN DE MARS 2017

LES SEULS DOCUMENT ADMIS SONT LES NOTES DE COURS MANUSCRITES. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE ET ORDINATEUR NON ADMIS.

Exercice 1

On suppose que les foies à greffer arrivent à l'hopital suivant un processus de Poisson N d'intensité $\lambda > 0$. Deux patients attendent une greffe de foie. Le premier a un temps de vie T_1 (avant greffe) qui suit une loi exponentielle de paramètre μ_1. Le second a un temps de vie T_2 (avant greffe) qui suit une loi exponentielle de paramètre μ_2. La politique de l'hopital est que le premier foie à greffer est greffé au premier patient si celui ci est encore en vie ; sinon, il est donné au second patient si celui ci est encore en vie. On suppose que N, T_1 et T_2 sont indépendants.

1. Calculer la probabilité que le premier patient est greffé.
2. Calculer la probabilité que le second patient est greffé.
3. Soit X le nombre de foies à greffer arrivés à l'hopital durant la période $[0, T_1]$. Calculer la fonction génératrice de X et en déduire que X suit une loi géométrique sur $\mathbb{N}$ dont on donnera explicitement le paramètre p.

Exercice 2

Un basketteur teste son habileté au lancer franc. Chaque panier augmente son score d'un unité mais un lancer raté lui fait perdre tous ses points. On note X_n son score à l'instant n. On modélise $(X_n)_{n\in\mathbb{N}}$ comme une chaîne de Markov à valeurs dans $\mathbb{N}$ de probabilités de transition

$$p(x, x+1) = \theta_x, \quad p(x, 0) = 1 - \theta_x$$

avec $\theta_x \in]0, 1[$ la probabilité de marquer le panier lorsqu'il a x points.

1. Soit $\tau_0 = \inf\{n \geq 1 : X_n = 0\}$. Calculer $u_n = \mathbb{P}_0(X_n = 0)$. En déduire $s_n = \mathbb{P}(\tau_0 \leq n)$.
2. En déduire que 0 est un état transient si et seulement si $\prod_{x=0}^{+\infty} \theta_x > 0$.
3. Etudier la récurrence et la transience de tous les autres états.
4. Calculer $m(x) = \mathbb{E}_0\left[\sum_{0\leq k<\tau_0} \mathbf{1}_{(X_k=x)}\right]$. Vérifier que m est bien une mesure invariante.

UNIVERSITÉ DE NANTES

E.C.N.

INTRODUCTION AUX PROCESSUS STOCHASTIQUES — ANNÉE 2018/2019

EXAMEN DE JANVIER 2019

LES SEULS DOCUMENTS ADMIS SONT LES NOTES DE COURS IMPRIMÉES. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE, TÉLÉPHONE ET ORDINATEUR NON ADMIS.

Exercice 1

On paramètre une droite D du plan par $(r,\theta) \in [0,+\infty[\times[0,2\pi[$ où $z = re^{i\theta}$ est la projection orthogonale de l'origine sur la droite D. On note $D = \phi(r,\theta)$.
On note $B_\rho = \{x \in \mathbb{C} : |x| \leq \rho\}$ la boule centrée en l'origine de rayon $\rho \geq 0$. On a ainsi

$$B_\rho \cap \phi(r,\theta) \neq \emptyset \iff r \leq \rho\,.$$

On considère sur $E = [0,+\infty[\times[0,2\pi[$ la mesure $\mu = \lambda \otimes \nu$ avec λ la mesure de Lebesgue et ν la probabilité uniforme, i.e. la mesure $d\nu(\theta) = \frac{1}{2\pi}\mathbf{1}_{(0\leq\theta<2\pi)}d\theta$.
Un réseau de droites aléatoires est fabriqué en considérant N une mesure aléatoire de Poisson sur E d'intensité μ.

1. Soit X_ρ la variable aléatoire qui compte le nombre de droites aléatoires qui intersectent la boule B_ρ. Donner la loi de X_ρ.
2. On considère qu'une boule de rayon ρ passe à travers le filtre aléatoire de droites si $X_\rho(\omega) = 0$. Calculer la probabilité qu'une boule de rayon ρ passe le filtre.
3. On considère Y une variable de loi exponentielle de paramètre $\alpha > 0$ indépendante de la mesure N. Calculer la probabilité qu'une boule de rayon Y passe le filtre.

Exercice 2

Une usine compte N machines et r mécaniciens (avec $r < N$). Chaque machine a une fiabilité p. On note X_n le nombre de machines en marche le jour n en début de journée. On suppose que chaque mécanicien peut réparer une machine défectueuse par jour.
On pose $\xi_i^{(n+1)} = 1$ si la i-ème machine qui marchait le matin du jour n est tombée en panne le jour n, et $\xi_i^{(n+1)} = 0$ sinon. Donc, sachant $X_n = x$, le nombre de machines tombées en panne est $Y_n = \sum_{i=1}^{x} \xi_i^{(n+1)}$. A la fin de la journée, chacun des r mécaniciens répare une machine si c'est nécessaire.

1. En déduire que l'on a la récurrence
$$X_{n+1} = X_n + \inf(r, N - X_n) - Y_n.$$
2. En déduire que X est une chaîne de Markov sur $E = \{0,\dots,N\}$.
3. Montrer que si $x \leq N - r$, et $r \leq z \leq r + x$ alors $x \to z$ (x mène à z). Montrer que si $x \geq N - r$ et $N - x \leq z \leq N$ alors $x \to z$.

4. En déduire que la chaîne est irréductible.
5. Montrer que la chaîne admet une unique probabilité invariante π.
6. Montrer que la chaîne est apériodique.
7. Montrer que $\mathbb{E}[X_n] \to \mu$ avec $\mu = \sum_i i\pi(i)$.
8. Dans le cas particulier $r = N$, donner une formule pour $\mathbb{E}[X_{n+1} \mid X_n]$ et en déduire μ .
9. Toujours dans ce cas particulier, montrer qu'il existe une unique probabilité réversible et retrouver la valeur de μ calculée précédemment.

Solution de l'Exercise .1

1. $X_\rho = N([0,\rho] \times [0,2\pi[) \sim P([0,\rho] \times [0,2\pi[) = P(\rho)$
2. $\mathbb{P}\big(X_\rho = 0\big) = e^{-\rho}$
3. Par Fubini positif

   $$\mathbb{P}(X_Y = 0) = \int \mathbb{P}\big(X_y = 0\big)\,dP_Y(y) = \int e^{-y}\,dP_Y(y) = \mathbb{E}\big[e^{-Y}\big] \tag{1}$$

   Pour $Y \sim \mathscr{E}(\alpha)$ il vient $\mathbb{P}(X_Y = 0) = \frac{\alpha}{\alpha+1}$.

Solution de l'Exercise .2

1. Au début de la journée n il y a donc $N - X_n$ machines en panne. On en répare le maximum possible avec r mécaniciens ie $\inf(r, N - X_n)$. Sachant que Y_n machines tombent en panne ce jour la, au début de la journée suivante il y a bien $X_n + \inf(r, N - X_n) - Y_n$ machines en marche.
2. La loi de Y_n sachant $X_n = x$ ne dépend pas de $X_{n-1}, X_{n-2}, \ldots$ et vaut $B(x,p)$.
3. Sachant $X_n = x$, Y_n prend avec une proba strictement positive toutes les valeurs entre 0 et x.
4. On obtient donc que N mène à tous les x. Réciproquement, x mène à $r + x$, $2r + x$, ... $z = kr + x$ avec $N - r \le z \le N$ et z mène à N. Donc la chaîne est irréductible
5. L'espace d'états est fini et la chaine irréductible.
6. $p_{N,N} > 0$ donc la chaîne est apériodique.
7. On a la convergence en loi de X_n vers $X_\infty \sim \pi$ En outre les va X_n sont bornées donc UI et $\mathbb{E}[X_n] \to \mathbb{E}[X_\infty] = \mu$.
8. Dans le cas particulier, $r = N$ on a $X_{n+1} = N - Y_n$ et donc

   $$\mathbb{E}[X_{n+1} \mid X_n] = N - pX_n$$

   Donc $\mathbb{E}[X_{n+1}] = N - p\mathbb{E}[X_n]$ et en passant à la limite, on obtient $\mu = \frac{N}{1+p}$.
9. On peut trouver une mesure réversible. $\pi(i) = Cp^{-i}\binom{N}{i}$ et on a $C = (1 + p^{-1})^N$.

UNIVERSITÉ DE NANTES **E.C.N.**

INTRODUCTION AUX PROCESSUS STOCHASTIQUES ANNÉE 2018/2019

EXAMEN DE MARS 2019

LES SEULS DOCUMENTS ADMIS SONT LES NOTES DE COURS IMPRIMÉES. TOUTES LES RÉPONSES DOIVENT ÊTRE SOIGNEUSEMENT JUSTIFIÉES. CALCULATRICE, TÉLÉPHONE ET ORDINATEUR NON ADMIS.

Exercice 1

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov irréductible sur un espace d'états fini I, de matrice de transition P.

1. Montrer que $(Y_n = (X_n, X_{n+1}), n \in \mathbb{N})$ est une chaîne de Markov $J = \{(i,j) : P(i,j) > 0\}$ de matrice de transition Q que l'on explicitera.
2. Montrer que $(Y_n)_{n\in\mathbb{N}}$ est irréductible.
3. Soit π la probabilité invariante pour P. Montrer que la probabilité invariante pour Q est $\lambda(i,j) = \pi(i)p_{ij}$.
4. On se fixe i, j. Montrer que la suite de variables
$$\hat{\theta}_n = \frac{\sum_{k=0}^n \mathbf{1}_{(X_k=i, X_{k+1}=j)}}{\sum_{k=0}^n \mathbf{1}_{(X_k=i)}}$$
converge presque sûrement vers p_{ij}.

Exercice 2

Soit $X_1, \dots, X_n$ des variables réelles IID, intégrables, et soit $S_k = X_1 + \cdots + X_k$ pour $1 \le k \le n$.

1. Soit f une fonction mesurable et $Z = f(X_1, X_2)$. Montrer que
$$\mathbb{E}[X_1 \mid Z] = \mathbb{E}[X_2 \mid Z].$$
2. En déduire que $\mathbb{E}[X_1 \mid S_2] = \frac{1}{2}S_2$.
3. Montrer de même que $\mathbb{E}[X_1 \mid S_k] = \frac{1}{k}S_k$.
4. Calculer $\mathbb{E}[S_k \mid S_n]$.
5. Soit $(\xi, \xi_i^{(n)}, n \ge 1, i \ge 1)$ une suite IID de variables entières de moyenne m, de loi ν. Le processus de Galton Watson de loi de reproduction ν issu de 1 est la suite de variables aléatoires donnée par la condition initiale $X_0 = 1$ et la récurrence
$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n+1)} \qquad (n \in \mathbb{N}).$$
Soit $(Y_n, n \in \mathbb{N})$ un autre processus de Galton Watson, issu de 1 lui aussi, de même loi de reproduction ν, indépendant de X, défini sur le même espace de probabilités, par exemple par la récurrence
$$Y_{n+1} = \sum_{X_n < i \le X_n + Y_n} \xi_i^{(n+1)}.$$
On suppose que $\mathbb{P}(\xi \ge 1) = 1$. Montrer que $Z_n = \frac{X_n}{X_n + Y_n}$ est une martingale.
6. Montrer que Z_n converge ps vers une variable aléatoire Z.
7. Calculer $\mathbb{E}[Z]$.

Solution de l'Exercise .1

1. On par la propriété de Makrov appliquée à l'instant $n+1$, si $\mathscr{F}_n = \sigma(X_k, k \leq n)$

   $$\mathbb{E}\left[f(Y_{n+1}) \mid \mathscr{F}_{n+1}\right] = \mathbb{E}_{f(X_0,X_1)\circ\theta^{n+1}|\mathscr{F}_{n+1}}[=]\mathbb{E}_{X_{n+1}}\left[f(X_0,X_1)\right] \tag{1}$$

   donc Y_n est une chaîne de Markov de noyau

   $$Qf(x,y) = \mathbb{E}_y\left[f(X_0,X_1)\right] = \mathbb{E}_y\left[f(y,X_1)\right] = \sum_t f(y,t)P(y,t)\,. \tag{2}$$

   En d'autres termes on a

   $$Q((x,y);(z,t)) = \mathbf{1}_{(y=z)}P(z,t)\,. \tag{3}$$

2. pour aller de (i,j) à (i',j') il suffit de trouver un chemin pour aller, pour la chaine X de j à i'.
3. C'est un simple calcul

   $$\lambda Q(k,l) = \sum_{(i,j)\in J} \lambda(i,j)Q((i,j),(k,l)) = \sum_{(i,j)\in J} \pi(i)P(i,j)\,\mathbf{1}_{(j=k)}P(k,l) = (\sum_i pi(i)P(i,k))P(k,l) = \pi(k)P(k,l) = \lambda(k,l)\,. \tag{4}$$