Sujet : Probabilités et Statistiques

Préambule

Ce sujet est consacré à l'étude de la marche persistante, qui est une chaîne de Markov $(X_n, V_n)_{n\in\mathbb{N}}$ sur $\mathbb{Z} \times \{-1, 1\}$ définie comme suit. On considère $U : \mathbb{Z} \to \mathbb{R}$ telle que

$$\mathcal{Z} := \sum_{k\in\mathbb{Z}} e^{-U(k)} < +\infty\,.$$

On cherche, par méthode de Monte-Carlo, à calculer des moyennes par rapport à la loi de probabilité π sur $\mathbb{Z}$ définie par

$$\pi(k) = \frac{1}{\mathcal{Z}} e^{-U(k)} \qquad \forall k \in \mathbb{Z}\,.$$

Pour cela, on considère $X_0 \in \mathbb{Z}$ et $V_0 \in \{-1, 1\}$ des conditions initiales et une suite $(A_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées (ce qu'on note dans la suite : i.i.d.) de loi $\mathcal{U}([0, 1])$ (uniforme sur $[0, 1]$), indépendante de (X_0, V_0), et on pose pour $n \in \mathbb{N}$

$$(X_{n+1}, V_{n+1}) = \begin{cases} (X_n + V_n, V_n) & \text{si } A_n \leqslant \min\left(1, e^{U(X_n)-U(X_n+V_n)}\right) \\ (X_n, -V_n) & \text{sinon.} \end{cases}$$

On appelle X_n et V_n respectivement la position et la vitesse de la marche persistante au temps n. Si $V_n = -V_{n-1}$, on dit qu'il y a une collision au temps n.

On rappelle que la loi géométrique de paramètre $p \in (0, 1)$, notée $\mathcal{G}(p)$, est la loi de probabilité sur $\mathbb{N}$ de fonction de répartition $F(k) = 1 - (1 - p)^k$ pour $k \in \mathbb{N}$.

Les 4 sections sont indépendantes. On pourra admettre le résultat d'une question pour l'utiliser par la suite.

1 Simulation de la chaîne

L'objectif de cette section est de proposer une méthode efficace pour simuler une trajectoire de la chaîne de Markov. Par “efficace”, on entend une méthode qui ne nécessite pas de calculer $U(X_n)$ à chaque temps $n \in \mathbb{N}$, ce qui peut s'avérer coûteux en pratique. On part du principe qu'un ordinateur peut générer des variables indépendantes de loi $\mathcal{U}([0, 1])$.

1. Soit A de loi $\mathcal{U}([0, 1])$ et $\alpha < 0$. Montrer que $\lceil \alpha \ln A \rceil$ est une variable aléatoire de loi géométrique dont on précisera le paramètre (on rappelle que, pour $x \in \mathbb{R}$, $\lceil x \rceil$ désigne la partie supérieure de x, c'est-à-dire le plus petit entier supérieur ou égal à x).
2. On note $T_1 = \inf\{n \in \mathbb{N}_*,\ V_n = -V_{n-1}\}$ le premier temps de collision. Pour tout $k \in \mathbb{N}$, $k < T_1$, exprimer (X_k, V_k) en fonction de (X_0, V_0) et k, et de même exprimer (X_{T_1}, V_{T_1}) en fonction de (X_0, V_0) et T_1.

L'objectif revient donc à simuler T_1 efficacement. Dans les questions suivantes, on suppose que U est Lipschitz, c'est-à-dire que $|U(x + 1) - U(x)| \leqslant L$ pour tout $x \in \mathbb{Z}$, avec une constante L connue. On pose $S_0 = 0$ et pour tout $k \in \mathbb{N}$,

$$S_{k+1} = \inf\{n > S_k,\ A_n > e^{-L}\}\,.$$

3. Pour tout $k \in \mathbb{N}$, quelle est la loi de $S_{k+1} - S_k$?

4. Montrer que, pour $i, j \in \mathbb{N}$ avec $i \neq j$, $S_{i+1} - S_i$ est indépendant de $S_{j+1} - S_j$.
5. Montrer que, pour $k \in \mathbb{N}_*$, la loi de S_k est donnée par

$$\mathbb{P}\left(S_k = s\right) = \binom{s-1}{k-1} e^{-L(s-k)}(1-e^{-L})^k \qquad \forall s \geqslant k$$

   (on pourra raisonner par récurrence ; ou bien remarquer qu'il y a autant de n-uplets d'entiers non nuls dont la somme vaut $m \in \mathbb{N}$ que de façons de mettre $n-1$ barrières entre m boules alignées, les entiers du n-uplet étant alors donnés par le nombre de boules entre deux barrières successives).

On considère une suite i.i.d. $(B_n)_{n\in\mathbb{N}}$ de loi $\mathcal{U}([0,1])$ indépendante de $(A_n)_{n\in\mathbb{N}}$ et de (X_0, V_0) et on pose

$$\tilde{T}_1 = \inf\left\{k \in \mathbb{N}_*,\ A_{k-1} > e^{-L} \text{ et } B_{k-1} \leqslant \frac{1 - e^{U(X_0+(k-1)V_0)-U(X_0+kV_0)}}{1-e^{-L}}\right\}.$$

6. Montrer que $\tilde{T}_1$ a la même loi que T_1.
7. Proposer une façon de simuler $(X_n, V_n)_{n\in\mathbb{N}}$ à partir de deux suites i.i.d. de variables aléatoires de loi uniformes sur $[0,1]$, de telle sorte que $U(X_k)$ ne soit pas systématiquement calculé pour tout $k \in \mathbb{N}$.

2 Invariance de la mesure de Gibbs

Soit $\mathcal{E}$ un ensemble discret, et $(Z_n)_{n\in\mathbb{N}}$ une chaîne de Markov à valeurs dans $\mathcal{E}$, qui est homogène, c'est-à-dire telle que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}\left(Z_{n+1} = z' \mid Z_n = z\right) = \mathbb{P}\left(Z_1 = z' \mid Z_0 = z\right).$$

On rappelle qu'à une telle chaîne est associé un noyau de transition $q : \mathcal{E}^2 \to [0,1]$ défini par

$$q\left(z, z'\right) = \mathbb{P}\left(Z_1 = z' \mid Z_0 = z\right) \qquad \forall z, z' \in \mathcal{E}.$$

Remarquons que, à $z \in \mathcal{E}$ fixé, $q(z, \cdot)$ est une loi de probabilité sur $\mathcal{E}$.

Un noyau de transition est dit symétrique si $q\left(z, z'\right) = q\left(z', z\right)$ pour tous $z, z' \in \mathcal{E}$. Pour ν une loi de probabilité sur $\mathcal{E}$, q est dit ν-réversible si

$$\nu(z) q\left(z, z'\right) = \nu(z') q\left(z', z\right) \qquad \forall z, z' \in \mathcal{E}.$$

D'autre part, ν est dite invariante pour $(Z_n)_{n\in\mathbb{N}}$ (ou, de façon équivalente, pour q) si le fait que Z_0 soit distribué selon ν entraîne que Z_1 est distribué selon ν.

1. Montrer que, si q est le noyau de transition associé à $(Z_n)_{n\in\mathbb{N}}$, alors ν est invariante pour $(Z_n)_{n\in\mathbb{N}}$ si et seulement si

$$\nu(z) \ = \ \sum_{z'\in\mathcal{E}} \nu(z') q(z', z) \qquad \forall z \in \mathcal{E}.$$

2. Montrer que, si ν est invariante pour $(Z_n)_{n\in\mathbb{N}}$ et si Z_0 est distribué selon ν, alors Z_n est distribué selon ν pour tout $n \in \mathbb{N}$.
3. Montrer que si q est ν-réversible, alors ν est invariante pour q.

Étant donnés deux noyaux de transition q_1 et q_2 sur $\mathcal{E}$, on définit une chaîne de Markov de la façon suivante. Pour tout $n \in \mathbb{N}$, étant donné $Y_n \in \mathcal{E}$, on tire aléatoirement $\tilde{Y}_n$ selon la loi $q_1(Y_n, \cdot)$, puis Y_{n+1} selon la loi $q_2(\tilde{Y}_n, \cdot)$. On note q_3 le noyau de transition de $(Y_n)_{n\in\mathbb{N}}$.

4. Exprimer q_3 en fonction de q_1 et q_2.
5. Montrer que, si ν est invariante pour q_1 et q_2, alors ν est invariante pour q_3.

Étant donné un noyau de transition q et une loi de probabilité ν sur $\mathcal{E}$, l'algorithme de Metropolis-Hastings définit une chaîne de Markov de la façon suivante. Pour tout $n \in \mathbb{N}$, étant donné $Y_n \in \mathcal{E}$ on tire aléatoirement $\tilde{Y}_n$ selon la loi $q\,(Y_n, \cdot)$ et $B_n \sim \mathcal{U}([0,1])$, et on pose

$$Y_{n+1} = \begin{cases} \tilde{Y}_n & \text{si } B_n \leq \min\left(1, \frac{\nu(\tilde{Y}_n)}{\nu(Y_n)}\right) \\ Y_n & \text{sinon.} \end{cases}$$

On note q_ν le noyau de transition de $(Y_n)_{n\in\mathbb{N}}$.

6. Exprimer q_ν en fonction de q et ν.
7. Montrer que, si q est symétrique, alors q_ν est ν-réversible.

Dans la suite de cette section, $\mathcal{E} = \mathbb{Z} \times \{-1, 1\}$. On considère une chaîne de Markov $(Y_n, W_n)_{n\in\mathbb{N}}$ sur $\mathcal{E}$ donnée pour tout $n \in \mathbb{N}$ par

$$(Y_{n+1}, W_{n+1}) = (Y_n + W_n, -W_n)\,.$$

8. Donner le noyau de transition de $(Y_n, W_n)_{n\in\mathbb{N}}$.
9. Montrer que ce dernier est symétrique.

On considère la loi de probabilité ν sur $\mathbb{Z} \times \{-1, 1\}$ donnée par

$$\nu(x, v) = \frac{1}{2}\pi(x) \qquad \forall x \in \mathbb{Z},\ v \in \{-1, 1\}\,,$$

et une chaîne de Markov $(Y'_n, W'_n)_{n\in\mathbb{N}}$ sur $\mathcal{E}$ donnée pour tout $n \in \mathbb{N}$ par

$$(Y'_{n+1}, W'_{n+1}) = (Y'_n, -W'_n)\,.$$

10. Montrer que le noyau de transition de $(Y'_n, W'_n)_{n\in\mathbb{N}}$ est ν-réversible.
11. Déduire de toutes les questions précédentes que ν est invariante pour la marche persistante $(X_n, V_n)_{n\in\mathbb{N}}$ définie dans le préambule du sujet.
12. Montrer que, si (X_0, V_0) est distribué selon la loi ν, alors X_n est distribué selon la loi π pour tout $n \in \mathbb{N}$.
13. Le noyau de transition de la marche persistante est-il réversible ?

3 Ergodicité dans un cas unimodal

Dans toute cette partie, on suppose que U est paire ($U(k) = U(-k)$ pour tout $k \in \mathbb{Z}$) et qu'il existe $r > 0$ tel que $U(k+1) - U(k) \geqslant r$ pour tout $k \geqslant 0$. En particulier, U est strictement décroissante sur $]\!]-\infty, 0]\!]$ et strictement croissante sur $[\![0, +\infty[\![$. On considère également une fonction $f : \mathbb{Z} \to \mathbb{R}$ qu'on suppose paire et bornée. L'objectif est d'étudier la limite quand n tend vers l'infini de la moyenne empirique

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

On suppose que $(X_0, V_0) = (0, 1)$ et l'on pose $R_0 = 0$ et, pour $n \in \mathbb{N}$,

$$\begin{array}{rcl} T_n & = & \inf\{k \geqslant R_n,\ V_{k+1} = -V_k\} - R_n \\ R_{n+1} & = & \inf\{k > R_n,\ X_k = 0\}\,. \end{array}$$

1. Supposons que, pour un certain $n \in \mathbb{N}$, $X_n V_n \leqslant 0$. Montrer qu'alors, presque sûrement, $X_{n+k} = X_n + kV_n$ pour tout $k \leqslant |X_n|$.
2. Montrer que pour tout $n \in \mathbb{N}$, $R_{n+1} - R_n = 2T_n + 1$.
3. Calculer V_{R_k} pour tout $k \in \mathbb{N}$.
4. Montrer que pour tout $k \in \mathbb{N}$,
$$\mathbb{P}\left(T_0 \geqslant k\right) = e^{U(0)-U(k)}\,.$$
5. En déduire que R_1 est d'espérance et de variance finie.

Pour tout $n \in \mathbb{N}$ on pose

$$Z_n = \sum_{k=R_n}^{R_{n+1}-1} f(X_k)\,.$$

On admettra que, du fait de la parité de U et f et de la propriété de Markov forte, les $\{Z_n,\ n \in \mathbb{N}\}$ sont indépendants et identiquement distribués.

6. Montrer que Z_0 est d'espérance et de variance finie.
7. Montrer que
$$\mathbb{E}\left(Z_0\right) = \lambda \sum_{k\in\mathbb{Z}} f(k)\pi(k) \qquad \text{avec } \lambda = \sum_{k\in\mathbb{Z}} e^{U(0)-U(k)}\,.$$
8. En déduire que $\mathbb{E}(R_{n+1} - R_n) = \lambda$ pour tout $n \in \mathbb{N}$.
9. On note $\mathcal{A}_n$ l'événement $\{|R_n/(n\lambda) - 1| \leqslant n^{-\frac{1}{4}}\}$. Montrer qu'il existe une constante $C > 0$ telle que pour tout $n \in \mathbb{N}_*$, $\mathbb{P}\left(\mathcal{A}_n\right) \geqslant 1 - C/\sqrt{n}$.
10. Pour $\varepsilon > 0$ et $n \in \mathbb{N}$ on note $\mathcal{B}_n^\varepsilon$ l'événement
$$\left\{\left|\frac{1}{n}\sum_{i=1}^n f(X_i) - \sum_{k\in\mathbb{Z}} f(k)\pi(k)\right| \geqslant \varepsilon\right\}\,.$$
Montrer que pour tout $\varepsilon > 0$, $\mathbb{P}\left(\mathcal{A}_{\lceil n/\lambda\rceil} \cap \mathcal{B}_n^\varepsilon\right)$ tend vers 0 quand n tend vers $+\infty$.
11. En déduire que $n^{-1}\sum_{i=1}^n f(X_i)$ converge en probabilité quand n tend vers l'infini vers une limite que l'on précisera.

4 Limite d'échelle

Dans cette partie, on considère $H \in \mathcal{C}^2(\mathbb{R})$ avec $\|H''\|_\infty < \infty$ et qui tend vers l'infini à l'infini, et $x \in \mathbb{R}$, $v \in \{-1, 1\}$ fixés. Pour tout $\delta > 0$, on considère $x_\delta \in \mathbb{Z}$ et on note $U_\delta : \mathbb{Z} \to \mathbb{R}$ la fonction définie par $U_\delta(k) = H(\delta k)$ pour $k \in \mathbb{Z}$. On considère la marche persistante $(X_n^\delta, V_n^\delta)_{n\in\mathbb{N}}$ associée à U_δ et avec conditions initiales (x_δ, v) telle que définie dans le préambule, c'est-à-dire qu'on pose $(X_0^\delta, V_0^\delta) = (x_\delta, v)$ et, pour tout $n \in \mathbb{N}$,

$$(X_{n+1}^\delta, V_{n+1}^\delta) = \left\{\begin{array}{ll} (X_n^\delta + V_n^\delta, V_n^\delta) & \text{si } A_n \leqslant \min\left(1, e^{U_\delta(X_n^\delta) - U_\delta(X_n^\delta + V_n^\delta)}\right) \\ (X_n^\delta, -V_n^\delta) & \text{sinon,} \end{array}\right.$$

avec $(A_n)_{n\in\mathbb{N}}$ une suite i.i.d. de variables aléatoires de loi $\mathcal{U}([0,1])$. Le but est d'étudier le comportement de la chaîne lorsque δ tend vers 0. On pose $T_0^\delta = 0$ et, pour tout $k \in \mathbb{N}_*$,

$$T_k^\delta = \inf\{n > T_{k-1},\ V_n^\delta = -V_{n-1}^\delta\}\,.$$

On s'intéresse d'abord au premier temps de collision T_1^δ.

1. Montrer que, pour $v \in \{-1,1\}$, $\int_0^t (vH'(vs))_+ \,\mathrm{d}s$ tend vers $+\infty$ quand t tend vers $+\infty$ (où, pour $y \in \mathbb{R}$, $(y)_+ := \max(0,y)$ désigne la partie positive de y).
2. Montrer que la fonction de répartition F_δ de T_1^δ est donnée par
$$F_\delta(k) = 1 - \exp\left(-\sum_{j=0}^{k-1}\Big(U_\delta(x_\delta + (j+1)v) - U_\delta(x_\delta + jv)\Big)_+\right) \qquad \forall k \in \mathbb{N}\,.$$
3. En déduire que, si δx_δ tend vers x quand δ tend vers 0, alors δT_1^δ converge en loi quand δ tend vers 0 vers une variable T_1^* dont on précisera la fonction de répartition F_* sur $\mathbb{R}$.
4. En déduire que, si δx_δ tend vers x quand δ tend vers 0, alors $(\delta X^\delta_{T_1^\delta}, V^\delta_{T_1^\delta})$ converge en loi quand δ tend vers 0 vers une variable aléatoire que l'on précisera (en l'exprimant en fonction de T_1^*).
5. Soit Y une variable aléatoire de fonction de répartition F. On rappelle que l'inverse généralisée F^{-1} de F est définie par $F^{-1}(u) = \inf\{t \in \mathbb{R},\ F(t) \geqslant u\}$ pour $u \in [0,1]$. Soit A une variable aléatoire de loi $\mathcal{U}([0,1])$. Montrer que $F^{-1}(A)$ a la même loi que Y.
6. Construire une variable aléatoire S_1^* de même loi que T_1^* et, pour tout $\delta > 0$, une variable aléatoire S_1^δ de même loi que T_1^δ, de telle sorte que, si δx_δ tend vers x quand δ tend vers 0, alors δS_1^δ converge presque sûrement vers S_1^* quand δ tend vers 0.

À partir de maintenant on suppose que $x_\delta = \lceil x/\delta \rceil$. Pour $\delta > 0$, $z \in \mathbb{Z}$ et $v \in \{-1,1\}$ on note $F_{\delta,z,v}$ la fonction de répartition sur $\mathbb{N}$ définie par

$$F_{\delta,z,v}(k) = 1 - \exp\left(-\sum_{j=0}^{k}\Big(U_\delta(z + (j+1)v) - U_\delta(z + jv)\Big)_+\right) \qquad \forall k \in \mathbb{N}\,.$$

On pose $(Y_0^\delta, W_0^\delta) = (x_\delta, v)$ puis, pour tout $n \in \mathbb{N}$,

$$S_n^\delta = F^{-1}_{\delta, Y_n^\delta, W_n^\delta}(A_n)\,, \qquad Y_{n+1}^\delta = Y_n^\delta + \big(S_n^\delta - 1\big)W_n^\delta \qquad \text{et} \qquad W_{n+1}^\delta = -W_n^\delta\,.$$

7. Montrer que pour tout $k \in \mathbb{N}$, $(Y_k^\delta, W_k^\delta, S_k^\delta)$ a la même loi que $(X^\delta_{T_k^\delta}, V^\delta_{T_k^\delta}, T^\delta_{k+1} - T_k^\delta)$.
8. Montrer que pour tout $N \in \mathbb{N}_*$, $(\delta Y_k^\delta, W_k^\delta, \delta S_k^\delta)_{k\in[\![0,N]\!]}$ converge presque sûrement quand δ tend vers 0 vers un vecteur aléatoire $(Y_k^*, W_k^*, S_k^*)_{k\in[\![0,N]\!]}$ dont on précisera la construction.

Pour tout $\delta > 0$, on définit $R_0^\delta = 0$ et, pour tout $n \geqslant 1$, $R_n^\delta = \sum_{k=0}^{n-1} S_k^\delta - 1/2$. Pour tout $n \in \mathbb{N}$ et tout $t \in [R_n^\delta, R_{n+1}^\delta[$, on pose $\tilde V_t^\delta = (-1)^n v$. Enfin, pour tout $t \geqslant 0$, on pose $\tilde X_t^\delta = x_\delta + \int_0^t \tilde V_s^\delta \mathrm{d}s$.

9. Montrer que $(\tilde X_n^\delta, \tilde V_n^\delta)_{n\in\mathbb{N}}$ a la même loi que $(X_n^\delta, V_n^\delta)_{n\in\mathbb{N}}$.

On pose $R_n^* = \sum_{k=0}^{n-1} S_k^*$ pour tout $n \in \mathbb{N}$. Pour tout $n \in \mathbb{N}$ et tout $t \in [R_n^*, R_{n+1}^*[$, on pose $\tilde V_t^* = (-1)^n v$. Enfin, pour tout $t \geqslant 0$, on pose $\tilde X_t^* = x + \int_0^t \tilde V_s^* \mathrm{d}s$.

10. Montrer que pour tout $N \in \mathbb{N}_*$ et pour tout $(t_1, \ldots, t_N) \in \mathbb{R}_*^N$,
$$\Big((\delta \tilde X^\delta_{t_1/\delta}, \tilde V^\delta_{t_1/\delta}), \ldots, (\delta \tilde X^\delta_{t_N/\delta}, \tilde V^\delta_{t_N/\delta})\Big) \underset{\delta\to 0}{\xrightarrow{p.s}} \Big((\tilde X^*_{t_1}, \tilde V^*_{t_1}), \ldots, (\tilde X^*_{t_N}, \tilde V^*_{t_N})\Big)$$
11. En déduire que pour tout $N \in \mathbb{N}_*$ et pour tout $(t_1, \ldots, t_N) \in \mathbb{R}_*^N$,
$$\big((\delta X^\delta_{\lceil t_1/\delta\rceil}, V^\delta_{\lceil t_1/\delta\rceil}), \ldots, (\delta X^\delta_{\lceil t_N/\delta\rceil}, V^\delta_{\lceil t_N/\delta\rceil})\big)$$
converge en loi, quand δ tend vers 0, vers une limite qu'on précisera.