Corrigé de l'examen du cours de M2 : Méthodes numériques probabilistes

15 janvier 2014, 09h00 - 12h00.

Les notes de cours sont autorisées.

Exercice : équations différentielles stochastiques et temps de sortie

1 On a évidemment
$$X_t^x = x + \alpha t + B_t.$$
La variable aléatoire X_t^x est gaussienne, centrée en $x + \alpha t$ et de variance t.

2.1 Il s'agit de la formule de Feynman-Kac vue en cours : $\forall x \in [a, b]$,
$$u(x) = \mathbb{E}\left(g 1_{\tau_a^x < \tau_b^x} + d 1_{\tau_b^x < \tau_a^x} - \int_0^{\tau^x} f(X_s^x)\, ds \right).$$

2.2 On choisit $\alpha = f = 0$ et $\beta = 1$. On a donc
$$u(x) = \mathbb{E}(1_{\tau_b^x < \tau_a^x}) = \mathbb{P}(\tau_b^x < \tau_a^x).$$
Par ailleurs, on peut résoudre facilement le problème du second ordre :
$$\begin{cases} \alpha u'(x) + \dfrac{1}{2} u''(x) = 0 \text{ pour tout } x \in (a, b) \\ u(a) = 0, \quad u(b) = 1. \end{cases}$$
On trouve $u(x) = \frac{e^{-2\alpha x} - e^{-2\alpha a}}{e^{-2\alpha b} - e^{-2\alpha a}}$ et donc, pour tout $x \in [a, b]$,
$$\mathbb{P}(\tau_b^x < \tau_a^x) = \frac{e^{-2\alpha x} - e^{-2\alpha a}}{e^{-2\alpha b} - e^{-2\alpha a}}.$$
Si $\alpha < 0$, on a $\lim_{a \to -\infty} \mathbb{P}(\tau_b^x < \tau_a^x) = \frac{e^{-2\alpha x}}{e^{-2\alpha b}}$. Si $\alpha > 0$, on a $\lim_{a \to -\infty} \mathbb{P}(\tau_b^x < \tau_a^x) = 1$. Donc, si $\alpha > 0$, presque sûrement, on atteint $b < \infty$ avant $-\infty$. En revanche, si $\alpha < 0$, on a une probabilité non nulle d'atteindre $-\infty$ avant $b < \infty$. Ceci est consistant avec le fait que le drift αt est positif et croît en temps quand $\alpha > 0$, et est négatif et décroît en temps quand $\alpha < 0$.

3 En supposant $Y_t^y > 0$, on a, par la formule d'Itô,
$$\begin{aligned} d \ln(Y_t^y) &= \frac{dY_t^y}{Y_t^y} - \frac{1}{2(Y_t^y)^2} (Y_t^y)^2\, dt \\ &= \left(r - \frac{1}{2} \right) dt + dB_t. \end{aligned}$$
On en déduit qu'une solution plausible pour l'EDS est
$$Y_t^y = y \exp\left(\left(r - \frac{1}{2} \right) t + B_t \right).$$

On vérifie à nouveau par un calcul d'Itô que ce processus est bien solution de l'EDS. Etant donnée que l'EDS admet une unique solution forte (cf. théorème du cours), c'est l'unique solution.

4 On remarque que presque sûrement $Y_t^y = \exp(X_t^x)$ si $\alpha = r - 1/2$ et $\exp(x) = y$. Par conséquent, si $\alpha = r - 1/2$, presque sûrement

$$\sigma_c^y = \tau_{\ln(c)}^{\ln(y)} \text{ et } \sigma_d^y = \tau_{\ln(d)}^{\ln(y)}.$$

On en déduit que

$$\begin{aligned}\mathbb{P}(\sigma_d^y < \sigma_c^y) &= \mathbb{P}\left(\tau_{\ln(d)}^{\ln(y)} < \tau_{\ln(c)}^{\ln(y)}\right)\\ &= \frac{y^{-2r+1} - c^{-2r+1}}{d^{-2r+1} - c^{-2r+1}}.\end{aligned}$$

Quand $r > 1/2$, on a $\lim_{c\to 0} \mathbb{P}(\sigma_d^y < \sigma_c^y) = 1$. Par conséquent, presque sûrement, le processus touche n'importe quel seuil $d > 0$ avant 0. Ceci montre que presque sûrement, le processus ne touche pas 0. Si $r < 1/2$, on a $\lim_{c\to 0} \mathbb{P}(\sigma_d^y < \sigma_c^y) = \frac{y^{-2r+1}}{d^{-2r+1}}$. Dans ce cas, le processus touche 0 avant $d > 0$ avec une probabilité positive égale à $1 - \frac{y^{-2r+1}}{d^{-2r+1}}$. Noter qu'une fois que le processus touche 0, il y reste pour tous les temps ultérieurs.

Problème : Simulation d'évènements rares

1 Par la loi forte de grands nombres, presque sûrement, $\lim_{n\to\infty} P_n = p$. On a

$$\mathrm{Var}(P_n) = \frac{\mathrm{Var}(1_{V(X^1)>a})}{n} = \frac{p(1-p)}{n}.$$

Par conséquent, $\frac{\sqrt{\mathrm{Var}(P_n)}}{p} = \sqrt{\frac{(1-p)}{np}} \sim \frac{1}{\sqrt{np}}$ tend vers l'infini quand $p \to 0$. Il faut avoir n de l'ordre de $1/p$ pour obtenir un résultat significatif. A nombre de tirages n fixé, cette méthode de Monte Carlo ne permet pas d'obtenir une estimée fiable.

2 On vérifie que

$$\begin{aligned}\mathbb{E}(\tilde{P}_n) &= \mathbb{E}\left(\frac{q(\tilde{X}^1)}{\tilde{q}(\tilde{X}^1)} 1_{V(\tilde{X}^1)>a}\right)\\ &= \int \frac{q(x)}{\tilde{q}(x)} 1_{V(x)>a} \tilde{q}(x)\, dx\\ &= \int q(x) 1_{V(x)>a}\, dx = p.\end{aligned}$$

Par la loi forte des grands nombres, presque sûrement, $\lim_{n\to\infty} \tilde{P}_n = \mathbb{E}(\tilde{P}_n) = p$. Par ailleurs,

$$\mathrm{Var}(P_n) = \frac{1}{n} \mathrm{Var}\left(\frac{q(\tilde{X}^1)}{\tilde{q}(\tilde{X}^1)} 1_{V(\tilde{X}^1)>a}\right)$$

et

$$\begin{aligned}\mathrm{Var}\left(\frac{q(\tilde X^1)}{\tilde q(\tilde X^1)}1_{V(\tilde X^1)>a}\right) &= \mathbb{E}\left(\frac{q(\tilde X^1)}{\tilde q(\tilde X^1)}1_{V(\tilde X^1)>a}\right)^2 - p^2\\ &= \int\left(\frac{q(x)}{\tilde q(x)}1_{V(x)>a}\right)^2\tilde q(x)\,dx - p^2\\ &= \int\frac{q^2(x)}{\tilde q(x)}1_{V(x)>a}\,dx - p^2.\end{aligned}$$

Bien sûr $\mathrm{Var}\left(\frac{q(\tilde X^1)}{\tilde q(\tilde X^1)}1_{V(\tilde X^1)>a}\right) \geq 0$ et on remarque qu'une variance nulle est obtenue pour le choix

$$\tilde q(x) = \frac{q(x)1_{V(x)>a}}{\int q(x)1_{V(x)>a}\,dx} = \frac{q(x)1_{V(x)>a}}{p}.$$

On a donc en principe une méthode d'échantillonnage d'importance exacte, si on sait simuler selon la loi $\tilde q(x)\,dx$, qui est la loi initiale de $X \sim q(x)\,dx$, conditionnée à $V(X) > a$. Bien sûr, en pratique, il est en général très délicat d'échantillonner la mesure conditionnelle $\tilde q(x)\,dx$. On peut utiliser une méthode de rejet mais très peu d'échantillons tomberont dans la zone d'intérêt $\{x,\, V(x) > a\}$, puisque p est petit. De plus, on ne connaît pas le facteur de normalisation $\int q(x)1_{V(x)>a}\,dx$ qui apparaît dans la formule définissant $\tilde q$ et donc dans $\tilde P_n$. En fait, ce facteur est exactement p, la probabilité qu'on veut calculer ! La méthode n'est donc pas implémentable en pratique.

3.1 Il s'agit d'une formule de conditionnement. On a en effet

$$\begin{aligned}p &= \mathbb{P}(V(X) > a_m)\\ &= \mathbb{P}(V(X) > a_m | V(X) > a_{m-1})\mathbb{P}(V(X) > a_{m-1})\\ &= \mathbb{P}(V(X) > a_m | V(X) > a_{m-1})\mathbb{P}(V(X) > a_{m-1} | V(X) > a_{m-2})\mathbb{P}(V(X) > a_{m-2})\\ &\vdots\\ &= \left(\prod_{j=1}^m \mathbb{P}(V(X) > a_j | V(X) > a_{j-1})\right)\mathbb{P}(V(X) > a_0)\end{aligned}$$

ce qui montre le résultat puisque $\mathbb{P}(V(X) > 0) = 1$.

3.2 Puisque les variables aléatoires $(X_j^i)_{i\geq 1, j\in\{1,\dots,m\}}$ sont indépendantes, on a

$$\begin{aligned}\mathbb{E}(\bar P_n) &= \prod_{j=1}^m \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n 1_{V(X_j^i)>a_j}\right) = \prod_{j=1}^m \mathbb{E}\left(1_{V(X_j^1)>a_j}\right)\\ &= \prod_{j=1}^m \int 1_{V(x)>a_j} q_{a_{j-1}}(x)\,dx = \prod_{j=1}^m \int 1_{V(x)>a_j}\frac{q(x)1_{V(x)>a_{j-1}}}{\int q(y)1_{V(y)>a_{j-1}}\,dy}\,dx\\ &= \prod_{j=1}^m \frac{\int 1_{V(x)>a_j} q(x)\,dx}{\int q(y)1_{V(y)>a_{j-1}}\,dy} = \prod_{j=1}^m \mathbb{P}(V(X) > a_j | V(X) > a_{j-1})\\ &= \prod_{j=1}^m p_j = p.\end{aligned}$$

Par ailleurs, par la loi forte des grands nombres, pour tout $j \in \{1, \ldots, m\}$, $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n 1_{V(X_j^i)>a_j} = \mathbb{E}\left(1_{V(X_j^1)>a_j}\right) = p_j$ p.s. et donc, p.s.

$$\lim_{n\to\infty} \bar{P}_n = \prod_{j=1}^m p_j = p.$$

De plus,

$$\begin{aligned}
\mathbb{E}\left(\left(\bar{P}_n\right)^2\right) &= \mathbb{E}\left(\left(\prod_{j=1}^m \frac{1}{n}\sum_{i=1}^n 1_{V(X_j^i)>a_j}\right)^2\right) = \prod_{j=1}^m \frac{1}{n^2}\mathbb{E}\left(\left(\sum_{i=1}^n 1_{V(X_j^i)>a_j}\right)^2\right) \\
&= \prod_{j=1}^m \frac{1}{n^2}\mathbb{E}\left(\sum_{i=1}^n \left(1_{V(X_j^i)>a_j}\right)^2 + 2\sum_{1\le i<k\le n} 1_{V(X_j^i)>a_j}1_{V(X_j^k)>a_j}\right) \\
&= \prod_{j=1}^m \frac{1}{n^2}\left(np_j + n(n-1)p_j^2\right) = \prod_{j=1}^m p_j^2\left(1+\frac{1}{n}\left(\frac{1}{p_j}-1\right)\right) \\
&= p^2\prod_{j=1}^m \left(1+\frac{1}{n}\left(\frac{1}{p_j}-1\right)\right).
\end{aligned}$$

Donc

$$\begin{aligned}
\mathrm{Var}(\bar{P}_n) &= p^2\prod_{j=1}^m \left(1+\frac{1}{n}\left(\frac{1}{p_j}-1\right)\right) - p^2 \\
&= p^2\left(1+\frac{1}{n}\sum_{j=1}^m\left(\frac{1}{p_j}-1\right) + O\left(\frac{1}{n^2}\right)\right) - p^2 \\
&= \frac{p^2}{n}\sum_{j=1}^m\left(\frac{1}{p_j}-1\right) + O\left(\frac{1}{n^2}\right).
\end{aligned}$$

En particulier, on a bien

$$\begin{aligned}
\lim_{n\to\infty} n\mathrm{Var}(\bar{P}_n) &= p^2\sum_{j=1}^m\left(\frac{1}{p_j}-1\right) \\
&= mp^2\left(-1+\frac{1}{m}\sum_{j=1}^m \frac{1}{p_j}\right).
\end{aligned}$$

3.3 La fonction ln étant concave, on a

$$\ln\left(\frac{1}{m}\sum_{j=1}^m \frac{1}{p_j}\right) \ge \frac{1}{m}\sum_{j=1}^m \ln\left(\frac{1}{p_j}\right) = \ln\left(\prod_{j=1}^m p_j^{-1/m}\right)$$

et donc

$$\frac{1}{m}\sum_{j=1}^m \frac{1}{p_j} \ge \left(\prod_{j=1}^m p_j\right)^{-1/m}.$$

En particulier,

$$\min_{\substack{(p_1,\ldots,p_m)\in(\mathbb{R}_+)^m \\ \text{tel que } \prod_{j=1}^m p_j = p}} \left(\frac{1}{m}\sum_{j=1}^m \frac{1}{p_j}\right) \geq \frac{1}{p^{1/m}},$$

et il est immédiat de vérifier que le minimum est atteint quand $p_1 = \ldots = p_m = p^{1/m}$. D'après la question précédente, minimiser la variance sur les niveaux revient à minimiser la moyenne $\frac{1}{m}\sum_{j=1}^m \frac{1}{p_j}$. Par conséquent, la variance est minimale quand les a_i sont tels que $p_1 = p_2 = \ldots = p_m = p^{1/m} = \alpha$. Dans ce cas, on a

$$\begin{aligned}\lim_{n\to\infty} n\mathrm{Var}(\bar{P}_n) &= mp^2\left(-1+\frac{1}{p^{1/m}}\right) \\ &= \frac{\ln p}{\ln\alpha}p^2\left(-1+\frac{1}{\alpha}\right).\end{aligned}$$

Soit la fonction $f : \alpha \in (0,1) \mapsto -\frac{1}{\ln\alpha}\left(-1+\frac{1}{\alpha}\right)$. On calcule

$$\begin{aligned}f'(\alpha) &= \frac{1}{\alpha(\ln\alpha)^2}\left(-1+\frac{1}{\alpha}\right)+\frac{1}{\alpha^2\ln\alpha} \\ &= \frac{1}{\alpha^2(\ln\alpha)^2}\left(-\alpha+1+\ln\alpha\right) \leq 0.\end{aligned}$$

La fonction f est décroissante, ce qui montre qu'on aimerait prendre (pour un p fixé) α le plus proche possible de 1, soit de manière équivalente m le plus grand possible.

3.4 Il y a essentiellement deux difficultés pratiques associées à cette méthode. Premièrement, on ne sait pas *a priori* comment choisir les niveaux a_i de manière à ce que les probabilités associées p_i soient égales. Deuxièmement, on ne sait pas *a priori* simuler selon les lois conditionnelles $q_{a_j}(x)\,dx$. L'objectif de la section suivante est de résoudre le premier problème (le choix des niveaux a_i) en implémentant une méthode qui va adapter les niveaux a_i de manière à ce que $p_1 = \ldots = p_m$.

4.1 La fonction de répartition F est croissante, donc Λ également. La fonction F est à valeurs dans $[0,1]$, donc Λ est positive. Par ailleurs, $\Lambda(0) = -\ln(1-F(0)) = 0$ puisque $F(0) = 0$ (cf. $V > 0$) et $\Lambda(a) = -\ln(1-F(a)) = -\ln p$ puisque $1-F(a) = 1-\mathbb{P}(V(X) \leq a) = \mathbb{P}(V(X) > a) = p$.

4.2 Soit $u \in (0,1]$. Soit x_n une suite minimisante associée à $F^{-1}(u)$: x_n converge par valeurs supérieures vers $F^{-1}(u)$ et $F(x_n) \geq u$. Comme F est continue, on a $F(x_n)$ converge vers $F(F^{-1}(u))$ et donc que $F(F^{-1}(u)) \geq u$. Soit $y_n < F^{-1}(u)$ une suite qui converge vers $F^{-1}(u)$ par valeurs inférieures. On a $F(y_n) < u$ (car si $F(y_n) \geq u$ alors $F^{-1}(u) \leq y_n$). Par conséquent, en passant à la limite, on obtient $F(F^{-1}(u)) \leq u$. On en déduit que $F(F^{-1}(u)) = u$.

Soit $U \sim \mathcal{U}(0,1)$. On a vu en cours (méthode de simulation par inversion de la fonction de répartition) que $F^{-1}(U)$ a même loi que $V(X)$. Donc $F(F^{-1}(U)) = U$ a même loi que $F(V(X))$. Ceci montre donc que $F(V(X)) \sim \mathcal{U}(0,1)$.

Pour montrer ensuite que $\Lambda(V(X)) \sim \mathcal{E}(1)$, on calcule, pour $z > 0$ (noter que $\Lambda(V(X)) \leq 0$),

$$\begin{aligned}
\mathbb{P}(\Lambda(V(X)) > z) &= \mathbb{P}(-\ln(1 - F(V(X))) > z) \\
&= \mathbb{P}(F(V(X)) > 1 - \exp(-z)) \\
&= \exp(-z)
\end{aligned}$$

ce qui montre bien que $\Lambda(V(X)) \sim \mathcal{E}(1)$.

Rappelons pourquoi, sans aucune hypothèse sur la fonction de répartition F, $F^{-1}(U)$ a même loi que $V(X)$. Ceci est basé sur l'équivalence,

$$\text{pour } x \in \mathbb{R}, \text{ et } u \in (0,1],\ F(x) \geq u \iff x \geq F^{-1}(u). \tag{1}$$

L'implication est triviale. Réciproquement, si $x \geq F^{-1}(u)$, comme F est croissante, on a $F(x) \geq F(F^{-1}(u))$. Soit alors x_n une suite minimisante associée à $F^{-1}(u)$: x_n converge par valeurs supérieures vers $F^{-1}(u)$ et $F(x_n) \geq u$. Comme F est continue à droite, on en déduit que $F(x_n)$ converge vers $F(F^{-1}(u))$ et donc que $F(F^{-1}(u)) \geq u$. Par conséquent $F(x) \geq u$.

Muni de cette équivalence, on a donc

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x).$$

Ceci implique que $F^{-1}(U)$ a même fonction de répartition que $V(X)$, donc que ces deux variables aléatoires ont même loi.

4.3 On a :

$$\begin{aligned}
\mathbb{P}(\Lambda(V(Y)) > z) &= \int 1_{\Lambda(V(y))>z} q_b(y)\, dy \\
&= \frac{\int 1_{\Lambda(V(y))>z} 1_{V(y)>b} q(y)\, dy}{\int 1_{V(y)>b} q(y)\, dy} \\
&= \frac{\mathbb{P}(\Lambda(V(X)) > \max(z, \Lambda(b)))}{\mathbb{P}(\Lambda(V(X)) > \Lambda(b))},
\end{aligned} \tag{2}$$

où X est de loi $q(x)\, dx$, et où nous avons utilisé le fait que Λ est une fonction croissante. On a également utilisé le fait que pour $X \sim q(x)\, dx$,

$$\mathbb{P}(V(X) > b) = \mathbb{P}(\Lambda(V(X)) > \Lambda(b))$$

Ceci revient à montrer que

$$\mathbb{P}(V(X) > b) = \mathbb{P}(F(V(X)) > F(b))$$

ce qui est équivalent à

$$\mathbb{P}(V(X) \leq b) = \mathbb{P}(F(V(X)) \leq F(b)).$$

En utilisant (1), on a

$$F(b) \geq F(V(X)) \iff b \geq F^{-1}(F(V(X))$$

Et donc

$$\begin{aligned}\mathbb{P}(F(b) \geq F(V(X))) &= \mathbb{P}(b \geq F^{-1}(F(V(X))) \\ &= \mathbb{P}(b \geq F^{-1}(U)) \\ &= \mathbb{P}(b \geq V(X)),\end{aligned}$$

ce qu'il fallait démontrer.

Pour terminer, on a montré à la question précédente que $\Lambda(V(X)) \sim \mathcal{E}(1)$. On déduit donc de (2) que

$$\mathbb{P}(\Lambda(V(Y)) > z) = \frac{\exp(-\max(z, \Lambda(b)))}{\exp(-\Lambda(b))}.$$

4.4 Si X et Y sont indépendantes, la loi du couple (X, Y) est une loi produit $\mu(dx)\nu(dy)$ (où μ désigne la loi de X et ν la loi de Y). On a donc par Fubini, pour f mesurable bornée,

$$\begin{aligned}\mathbb{E}(f(X,Y)) &= \int\int f(x,y)\mu(dx)\nu(dy) \\ &= \int \left(\int f(x,y)\nu(dy)\right)\mu(dx) \\ &= \mathbb{E}(g(X)),\end{aligned}$$

où $g(x) = \int f(x,y)\nu(dy) = \mathbb{E}(f(x,Y))$.

4.5 Pour $z_1, \ldots, z_n, z$ des nombres positifs (les variables aléatoires considérées ayant pour support $\mathbb{R}_+$), on considère

$$\begin{aligned}&\mathbb{P}(\Lambda(V(X_1^1)) - \Lambda(A_0) > z_1, \ldots, \Lambda(V(X_1^n)) - \Lambda(A_0) > z_n, \Lambda(A_0) > z) \\ &= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_1^1)) - \Lambda(A_0) > z_1, \ldots, \Lambda(V(X_1^n)) - \Lambda(A_0) > z_n, \Lambda(A_0) > z, I_0 = i_0) \\ &= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(X_0^{i_0})) > z_1, \ldots, \Lambda(V(X_1^{i_0})) - \Lambda(V(X_0^{i_0})) > z_{i_0}, \ldots, \Lambda(V(X_0^n)) - \Lambda(V(X_0^{i_0})) > z_n, \\ &\qquad \Lambda(V(X_0^{i_0})) > z, I_0 = i_0),\end{aligned}$$

où $X_1^{i_0}$ a pour loi $q_{A_0}(x)\,dx$. Dans la dernière égalité, on a utilisé le fait que si $I_0 = i_0$, alors $A_0 = V(X_0^{I_0})$, et $X_1^i = X_0^i$ pour $i \neq i_0$. Puisque les z_i sont positifs, on peut oublier la contrainte $I_0 = i_0$ de sorte que

$$\begin{aligned}&\mathbb{P}(\Lambda(V(X_1^1)) - \Lambda(A_0) > z_1, \ldots, \Lambda(V(X_1^n)) - \Lambda(A_0) > z_n, \Lambda(A_0) > z) \\ &= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(X_0^{i_0})) > z_1, \ldots, \Lambda(V(X_1^{i_0})) - \Lambda(V(X_0^{i_0})) > z_{i_0}, \ldots, \Lambda(V(X_0^n)) - \Lambda(V(X_0^{i_0})) > z_n, \\ &\qquad \Lambda(V(X_0^{i_0})) > z).\end{aligned}$$

Noter que les variables aléatoires $X_0^1, \dots, X_0^n, X_1^{i_0}$ sont indépendantes. On a donc, en conditionnant par $V(X_0^{i_0})$ (cf. la question précédente),

$$\mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(X_0^{i_0})) > z_1, \dots, \Lambda(V(X_1^{i_0})) - \Lambda(V(X_0^{i_0})) > z_{i_0}, \dots, \Lambda(V(X_0^n)) - \Lambda(V(X_0^{i_0})) > z_n,$$
$$\Lambda(V(X_0^{i_0})) > z) = \mathbb{E}\left(f(X_0^{i_0}) 1_{\Lambda(V(X_0^{i_0}))>z}\right)$$

avec, en utilisant à nouveau l'indépendance

$$\begin{aligned} f(x) &= \mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(x)) > z_1, \dots, \Lambda(V(X_1^{i_0})) - \Lambda(V(x)) > z_{i_0}, \dots, \Lambda(V(X_0^n)) - \Lambda(V(x)) > z_n) \\ &= \mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(x)) > z_1) \dots \mathbb{P}(\Lambda(V(X_1^{i_0})) - \Lambda(V(x)) > z_{i_0}) \dots \mathbb{P}(\Lambda(V(X_0^n)) - \Lambda(V(x)) > z_n). \end{aligned}$$

En utilisant la question 4.2, pour $i \neq i_0$, on a $\mathbb{P}(\Lambda(V(X_0^i)) - \Lambda(V(x)) > z_i) = \exp(-z_i - \Lambda(V(x)))$. En utilisant la question 4.3, pour $i = i_0$

$$\begin{aligned} \mathbb{P}(\Lambda(V(X_1^{i_0})) - \Lambda(V(x)) > z_{i_0}) &= \exp(\Lambda(V(x)) - \max(z_{i_0} + \Lambda(V(x)), \Lambda(V(x)))) \\ &= \exp(-z_{i_0}). \end{aligned}$$

On obtient donc, en utilisant le fait que $\Lambda(V(X_0^{i_0})) \sim \mathcal{E}(1)$,

$$\mathbb{P}(\Lambda(V(X_0^1)) - \Lambda(V(X_0^{i_0})) > z_1, \dots, \Lambda(V(X_1^{i_0})) - \Lambda(V(X_0^{i_0})) > z_{i_0}, \dots, \Lambda(V(X_0^n)) - \Lambda(V(X_0^{i_0})) > z_n,$$
$$\begin{aligned} \Lambda(V(X_0^{i_0})) > z) &= \mathbb{E}\left(\exp(-(z_1 + \ldots + z_n)) \exp(-(n-1)\Lambda(V(X_0^{i_0}))) 1_{\Lambda(V(X_0^{i_0}))>z}\right) \\ &= \exp(-(z_1 + \ldots + z_n)) \int_z^\infty \exp(-(n-1)y) \exp(-y)\, dy \\ &= \frac{1}{n} \exp(-(z_1 + \ldots + z_n)) \exp(-nz). \end{aligned}$$

Noter que cette expression ne dépend pas de i_0. Finalement,

$$\mathbb{P}(\Lambda(V(X_1^1)) - \Lambda(A_0) > z_1, \dots, \Lambda(V(X_1^n)) - \Lambda(A_0) > z_n, \Lambda(A_0) > z) = \exp(-(z_1 + \ldots + z_n)) \exp(-nz),$$

ce qui prouve le résultat. Noter qu'en conséquence $\Lambda(A_1) - \Lambda(A_0)$ est indépendant de $\Lambda(A_0)$ puisque

$$\Lambda(A_1) - \Lambda(A_0) = \min_{i \in \{1,\dots,n\}} \Lambda(V(X_1^i)) - \Lambda(A_0)$$

et les variables aléatoires $(\Lambda(V(X_1^i)) - \Lambda(A_0))_{1 \leq i \leq n}$ sont indépendantes de $\Lambda(A_0)$.

4.6 Soit $j \geq 2$. En raisonnant par récurrence, on suppose la propriété vérifiée pour $j-1$ et on veut la montrer pour j. Il faut donc montrer que $(\Lambda(V(X_j^i)) - \Lambda(A_{j-1}))_{1 \leq i \leq n}$ sont i.i.d. de loi exponentielle de paramètre 1, et sont indépendantes de $(\Lambda(A_{k-1}) - \Lambda(A_{k-2}))_{1 \leq k \leq j}$. Le raisonnement est similaire au cas $j = 1$.

On considère, pour $z_1, \dots, z_n, t_1, \dots, t_j$ des nombres positifs (les variables aléatoires considérées ayant pour support $\mathbb{R}_+$) :

$$\mathbb{P}(\Lambda(V(X_j^1)) - \Lambda(A_{j-1}) > z_1, \ldots, \Lambda(V(X_j^n)) - \Lambda(A_{j-1}) > z_n, \Lambda(A_0) > t_1, \ldots, \Lambda(A_{j-1}) - \Lambda(A_{j-2}) > t_j)$$
$$= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_j^1)) - \Lambda(A_{j-1}) > z_1, \ldots, \Lambda(V(X_j^n)) - \Lambda(A_{j-1}) > z_n, \Lambda(A_0) > t_1, \ldots, \Lambda(A_{j-1}) - \Lambda(A_{j-2}) > t_j,$$
$$I_{j-1} = i_0)$$

$$= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_{j-1}^1)) - \Lambda(V(X_{j-1}^{i_0})) > z_1, \ldots, \Lambda(V(X_j^{i_0})) - \Lambda(V(X_{j-1}^{i_0})) > z_{i_0}, \ldots,$$
$$\Lambda(V(X_{j-1}^n)) - \Lambda(V(X_{j-1}^{i_0})) > z_n, \Lambda(A_0) > t_1, \ldots, \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j, I_{j-1} = i_0)$$

$$= \sum_{i_0=1}^n \mathbb{P}(\Lambda(V(X_{j-1}^1)) - \Lambda(A_{j-2}) > z_1 + \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}), \ldots, \Lambda(V(X_j^{i_0})) - \Lambda(V(X_{j-1}^{i_0})) > z_{i_0}, \ldots,$$
$$\Lambda(V(X_{j-1}^n)) - \Lambda(A_{j-2}) > z_n + \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}), \Lambda(A_0) > t_1, \ldots, \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j)$$

En conditionnant par $\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}), \Lambda(A_{j-2}) - \Lambda(A_{j-3}), \ldots, \Lambda(A_0)$, et en utilisant l'hypothèse de récurrence au rang $j-1$, on a alors :

$$\mathbb{P}(\Lambda(V(X_{j-1}^1)) - \Lambda(A_{j-2}) > z_1 + \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}), \ldots, \Lambda(V(X_j^{i_0})) - \Lambda(V(X_{j-1}^{i_0})) > z_{i_0}, \ldots,$$
$$\Lambda(V(X_{j-1}^n)) - \Lambda(A_{j-2}) > z_n + \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}), \Lambda(A_0) > t_1, \ldots, \Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j)$$
$$= \mathbb{E}(\exp(-z_1 - \Lambda(V(X_{j-1}^{i_0})) + \Lambda(A_{j-2})) \ldots \exp(-z_{i_0}) \ldots \exp(-z_n - \Lambda(V(X_{j-1}^n)) + \Lambda(A_{j-2}))$$
$$1_{\Lambda(A_0) > t_1}, \ldots, 1_{\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j})$$
$$= \exp(-(z_1 + \ldots + z_n)) \mathbb{E}\Big(\exp\Big(-(n-1)(\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}))\Big) 1_{\Lambda(A_0) > t_1}, \ldots, 1_{\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j}\Big)$$
$$= \exp(-(z_1 + \ldots + z_n)) \mathbb{E}\Big(\exp\Big(-(n-1)(\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}))\Big) 1_{\Lambda(V(X_{j-1}^{i_0})) - \Lambda(A_{j-2}) > t_j}\Big)$$
$$\mathbb{P}(\Lambda(A_0) > t_1) \ldots \mathbb{P}(\Lambda(A_{j-2}) - \Lambda(A_{j-3}) > t_{j-1})$$
$$= \exp(-(z_1 + \ldots + z_n)) \frac{1}{n} \exp(-nt_j) \exp(-nt_1) \ldots \exp(-nt_{j-1}).$$

Noter que cette expression ne dépend pas de i_0. On obtient donc :

$$\mathbb{P}(\Lambda(V(X_j^1)) - \Lambda(A_{j-1}) > z_1, \ldots, \Lambda(V(X_j^n)) - \Lambda(A_{j-1}) > z_n, \Lambda(A_0) > t_1, \ldots, \Lambda(A_{j-1}) - \Lambda(A_{j-2}) > t_j)$$
$$= \exp(-(z_1 + \ldots + z_n)) \exp(-n(t_1 + \ldots t_j)).$$

Ceci montre la propriété de récurrence au rang j.

4.7 D'après les questions précédentes (noter que $\Lambda(A_{-1}) = \Lambda(0) = 0$),

$$\Lambda(A_j) = \sum_{k=0}^j \Lambda(A_k) - \Lambda(A_{k-1})$$

est une somme de variables aléatoires indépendantes de loi $\mathcal{E}(n)$. Par conséquent, si on note $(E_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. de loi $\mathcal{E}(n)$ (et en utilisant le fait $\Lambda(a) = -\ln p$),

$$\begin{aligned}
\mathbb{P}(\hat{J}_n = j) &= \mathbb{P}(A_j > a, A_{j-1} \leq a) \\
&= \mathbb{P}\left(\sum_{k=0}^{j} E_k > \Lambda(a), \sum_{k=0}^{j-1} E_k \leq \Lambda(a)\right) \\
&= \mathbb{P}\left(E_j > -\ln p - \sum_{k=0}^{j-1} E_k, \sum_{k=0}^{j-1} E_k \leq -\ln p\right) \\
&= \mathbb{E}\left(\exp\left(n\left(\ln p + \sum_{k=0}^{j-1} E_k\right)\right) 1_{\sum_{k=0}^{j-1} E_k \leq -\ln p}\right) \\
&= p^n n^j \int_0^\infty \dots \int_0^\infty 1_{\sum_{k=0}^{j-1} x_k \leq -\ln p} \, dx_0 \dots dx_{j-1}.
\end{aligned}$$

Une récurrence simple permet de calculer :

$$\int_0^\infty \dots \int_0^\infty 1_{\sum_{k=0}^{j-1} x_k \leq \alpha} \, dx_0 \dots dx_{j-1} = \frac{\alpha^j}{j!}.$$

On a donc finalement

$$\mathbb{P}(\hat{J}_n = j) = p^n n^j \frac{(-\ln p)^j}{j!}.$$

On calcule ensuite :

$$\begin{aligned}
\mathbb{E}(\hat{P}_n) &= \mathbb{E}\left(\left(1 - \frac{1}{n}\right)^{\hat{J}_n}\right) \\
&= \sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^j \mathbb{P}(\hat{J}_n = j) = \sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^j p^n n^j \frac{(-\ln p)^j}{j!} \\
&= p^n \sum_{j=0}^{\infty} \frac{((1-n)\ln p)^j}{j!} = p^n \exp((1-n)\ln p) = p.
\end{aligned}$$

4.8 On calcule

$$\begin{aligned}
\mathbb{E}((\hat{P}_n)^2) &= \mathbb{E}\left(\left(1 - \frac{1}{n}\right)^{2\hat{J}_n}\right) \\
&= \sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^{2j} \mathbb{P}(\hat{J}_n = j) = \sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^{2j} p^n n^j \frac{(-\ln p)^j}{j!} \\
&= p^n \sum_{j=0}^{\infty} \frac{-((1-n)^2(\ln p)/n)^j}{j!} = p^n \exp(-(1-n)^2(\ln p)/n) = p^{-1/n+2}.
\end{aligned}$$

On a donc

$$\mathrm{Var}(\hat{P}_n) = p^2 \left(p^{-1/n} - 1\right).$$

Dans la limite $p \to 0$ on a donc

$$\begin{aligned}\frac{\sqrt{\operatorname{Var}(\hat{P}_n)}}{p} &= \left(p^{-1/n} - 1\right)^{1/2} \\ &\sim p^{-1/(2n)}\end{aligned}$$

qui est beaucoup plus petite que l'erreur relative en $(np)^{-1/2}$ obtenue avec la méthode de Monte Carlo naïve de la première question.

Pour mettre en oeuvre l'algorithme, il reste à savoir simuler les lois conditionnelles $q_{A_j}(x)\,dx$. Pour cela, on peut utiliser un algorithme de Metropolis Hastings. On renvoie aux articles [F. Cérou et A. Guyader, Stoch. Anal. Appl. 25, 417 (2007)] [A. Guyader, N. Hengartner et E. Matzner-Løber, Appl. Math. Optim. 64, 171 (2011)] pour plus de précisions.