Examen du cours de M2 :
Méthodes numériques probabilistes

16 janvier 2015, 13h30 - 16h30.

Les notes de cours sont autorisées.

Exercice 1 : Stratification

Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. On suppose que l'on sait simuler des réalisations indépendantes de X. On veut calculer une approximation de $I = \mathbb{E}(f(X))$ par une méthode de Monte Carlo, où $f : \mathbb{R}^d \to \mathbb{R}$ est une fonction telle que $\mathbb{E}(|f(X)|^2) < \infty$.

1 Rappeler comment on peut approcher I par une moyenne empirique utilisant n réalisations indépendantes de X. Donner une expression de l'erreur dans la limite $n \to \infty$.

2 On suppose désormais que l'on dispose d'une partition de $\mathbb{R}^d$ en m sous-ensembles (appelés strates) :

$$\mathbb{R}^d = \bigcup_{i=1}^{m} A_i \text{ et } \forall i, j \in \{1, \dots, m\}, A_i \cap A_j = \emptyset.$$

On suppose que l'on connaît les proportions (supposées strictement positives) de chacune des strates pour la loi de X :

$$p_i = \mathbb{P}(X \in A_i) > 0$$

et que l'on sait simuler des réalisations indépendantes selon la loi conditionnelle de X sachant que $X \in A_i$. Pour $i \in \{1, \dots m\}$, on considère des réalisations $(X_k^i)_{k\geq 1}$ selon la loi conditionnelle de X sachant que $X \in A_i$, telles que les variables aléatoires $(X_k^i)_{i\in\{1,\dots,m\},\, k\geq 1}$ sont indépendantes. Pour m entiers non nuls $(n_1, \dots, n_m)$, on considère

$$\hat{I} = \sum_{i=1}^{m} p_i \hat{I}_i$$

où

$$\hat{I}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} f(X_k^i).$$

On note $\mu_i = \mathbb{E}(f(X_1^i)) = \mathbb{E}(f(X)|X \in A_i)$ et $\sigma_i^2 = \mathrm{Var}(f(X_1^i)) = \mathbb{E}(f^2(X)|X \in A_i) - (\mathbb{E}(f(X)|X \in A_i))^2$.

2.1 Vérifier que

$$\mathrm{Var}(f(X)) = \sum_{i=1}^{m} p_i \sigma_i^2 + \sum_{i=1}^{m} p_i \left(\mu_i - \sum_{j=1}^{m} p_j \mu_j \right)^2.$$

Donner une interprétation des deux termes au membre de droite.

2.2 Calculer $\mathbb{E}(\hat{I})$. Quel est le comportement de $\hat{I}$ quand $\min(n_1, \dots, n_m) \to \infty$?

2.3 Montrer que
$$\mathrm{Var}(\hat{I}) = \sum_{i=1}^{m} \frac{p_i^2 \sigma_i^2}{n_i}.$$

3 On suppose que le nombre total d'échantillons $n = \sum_{i=1}^{m} n_i$ est fixé, et on cherche une bonne allocation $(n_i)_{i\in\{1,\ldots,m\}}$.

3.1 Montrer que
$$\left(\sum_{i=1}^{m} p_i \sigma_i\right)^2 \leq n \sum_{i=1}^{m} \frac{p_i^2 \sigma_i^2}{n_i}.$$
En déduire que la variance de $\hat{I}$ est minimisée pour le choix (allocation optimale)
$$n_i^\star = n \, \frac{p_i \sigma_i}{\sum_{j=1}^{m} p_j \sigma_j}$$
où on suppose pour simplifier que pour tout $i \in \{1,\ldots,m\}$, $n \, \frac{p_i \sigma_i}{\sum_{j=1}^{m} p_j \sigma_j} \in \mathbb{N}$. Pourquoi est-ce que cette allocation est rarement utilisable en pratique ?

3.2 On choisit dans cette question (allocation proportionnelle)
$$n_i = np_i$$
en supposant pour simplifier que pour tout $i \in \{1,\ldots,m\}$, $np_i \in \mathbb{N}$. Montrer que dans ce cas
$$n\mathrm{Var}(\hat{I}) \leq \mathrm{Var}(f(X)).$$
Interpréter ce résultat. Enoncer un théorème central limite dans la limite $n \to \infty$. Comment faut-il choisir les strates pour diminuer l'erreur statistique ?

Exercice 2 : Variable de contrôle

Soit X une variable aléatoire réelle de carré intégrable. On veut calculer une approximation de $I = \mathbb{E}(X)$ par une méthode de Monte Carlo. On considère pour cela une variable aléatoire réelle de carré intégrable Y telle que $\mathbb{E}(Y) = 0$ et on considère l'estimateur, pour un réel α à choisir,
$$\hat{I} = \frac{1}{n} \sum_{k=1}^{n} (X_k - \alpha Y_k)$$
où $(X_k, Y_k)_{k\geq 1}$ désignent des réalisations indépendantes du couple (X, Y).

1 Quel est le comportement de $\hat{I}$ dans la limite $n \to \infty$? Calculer $\mathbb{E}(\hat{I})$ et $\mathrm{Var}(\hat{I})$. Montrer que la variance est minimisée pour
$$\alpha^\star = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}.$$
(On rappelle que $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.) Discuter dans quel cas la variance minimale est nulle. Dans ce cas, on dit que Y est une variable de contrôle idéale pour X.

2 En pratique, $\alpha^\star$ n'est pas connu. On propose donc de l'estimer par une méthode de Monte Carlo et donc de considérer

$$\tilde{I} = \frac{1}{n}\sum_{k=1}^{n}(X_k - \alpha_n Y_k)$$

où

$$\alpha_n = \frac{\frac{1}{n}\sum_{k=1}^n X_k Y_k - \left(\frac{1}{n}\sum_{k=1}^n X_k\right)\left(\frac{1}{n}\sum_{k=1}^n Y_k\right)}{\frac{1}{n}\sum_{k=1}^n (Y_k)^2 - \left(\frac{1}{n}\sum_{k=1}^n Y_k\right)^2}.$$

Que dire de la limite de $\tilde{I}$ quand $n \to \infty$? Enoncer un théorème central limite pour $\tilde{I}$.

3 On suppose désormais que $X = \varphi(Z_T)$ où $\varphi : \mathbb{R} \to \mathbb{R}$ est une fonction bornée et $(Z_t)_{0\le t\le T}$ est un processus stochastique à valeurs réelles tel que : pour tout $t \in [0, T]$,

$$Z_t = z_0 + \int_0^t b(s, Z_s)\, ds + \int_0^t \sigma(s, Z_s)\, dW_s$$

où la condition initiale $z_0 \in \mathbb{R}$ est déterministe, b et σ sont des fonctions à valeurs réelles régulières bornées, et $(W_t)_{t\ge 0}$ est un mouvement brownien. On suppose que b, σ et φ sont tels qu'il existe une fonction $u : (t, z) \in [0, T] \times \mathbb{R} \mapsto u(t, z) \in \mathbb{R}$ régulière vérifiant : $\|\partial_z u\|_{L^\infty([0,T]\times\mathbb{R})} < \infty$ et

$$\begin{cases} \partial_t u(t, z) + b(t, z)\, \partial_z u(t, z) + \dfrac{\sigma(t, z)^2}{2}\partial_{z,z} u(t, z) = 0, \\ \hfill u(T, z) = \varphi(z). \end{cases}$$

On pose $Y = \int_0^T \partial_z u(s, Z_s)\sigma(s, Z_s) dW_s$.

3.1 Vérifier que $\mathbb{E}(Y) = 0$ et montrer que $\mathrm{Var}(X - Y) = 0$ (Indication : calculer $d(u(t, Z_t))$). En déduire que Y est une variable de contrôle idéale pour X. Interpréter ce résultat.

3.2 Comment généraliser le résultat de la question précédente pour construire une variable de contrôle idéale pour $X = \exp\left(\int_0^T \psi(Z_s)\, ds\right)\varphi(Z_T)$, où ψ est une fonction à valeurs réelles bornée ?

Exercice 3 : Transformation de Girsanov

Soit X_t un processus à valeurs réelles tel que :

$$X_t = -h(t) + W_t$$

où $h : \mathbb{R}_+ \to \mathbb{R}$ est une fonction (déterministe) de classe $\mathcal{C}^2$ telle que $h(0) = 0$ et $(W_t)_{t\ge 0}$ est un mouvement brownien. On note $(\mathcal{F}_t)_{t\ge 0}$ la filtration engendrée par $(W_t)_{t\ge 0}$.

1 Soit $t > 0$ et $n \in \mathbb{N}$ deux constantes fixées, et $0 = t_0 < t_1 < \ldots < t_n = t$ une subdivision de l'intervalle $[0, t]$. Montrer que le vecteur $(X_{t_1}, \ldots, X_{t_n})$ a pour densité : pour $(x_1, \ldots, x_n) \in \mathbb{R}^n$ (avec $x_0 = 0$ par convention),

$$\left(\prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}}\right) \exp\left(-\sum_{i=0}^{n-1} \frac{(x_{i+1} - x_i + h(t_{i+1}) - h(t_i))^2}{2(t_{i+1} - t_i)}\right).$$

2 Montrer que pour toute fonction test bornée $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}\left(\varphi(X_{t_1}, \dots, X_{t_n})\right) = \mathbb{E}\left(\varphi(W_{t_1}, \dots, W_{t_n}) Z_t^{\mathrm{disc}}\right) \tag{1}$$

où

$$Z_t^{\mathrm{disc}} = \exp\left(-\sum_{i=0}^{n-1} \frac{h(t_{i+1}) - h(t_i)}{t_{i+1} - t_i}(W_{t_{i+1}} - W_{t_i}) - \frac{1}{2}\sum_{i=0}^{n-1} \frac{(h(t_{i+1}) - h(t_i))^2}{t_{i+1} - t_i}\right).$$

3 Dans la limite $n \to \infty$ et $\max_{i \in \{1,\dots,n\}}(t_{i+1} - t_i) \to 0$, on peut montrer que Z_t^{disc} converge vers

$$Z_t = \exp\left(-\int_0^t h'(s) dW_s - \frac{1}{2}\int_0^t (h'(s))^2\, ds\right)$$

et la relation (1) devient : pour toute fonction test bornée $\varphi : \mathcal{C}^0((0,t), \mathbb{R}) \to \mathbb{R}$,

$$\mathbb{E}\left(\varphi((X_s)_{s \in (0,t)})\right) = \mathbb{E}\left(\varphi((W_s)_{s \in (0,t)}) Z_t\right). \tag{2}$$

Rappeler pourquoi l'intégrale stochastique dans Z_t est bien définie et montrer que

$$dZ_t = -h'(t) Z_t\, dW_t.$$

4 Montrer que

$$Z_t = \exp\left(-h'(t) W_t + \int_0^t h''(s) W_s\, ds - \frac{1}{2}\int_0^t (h'(s))^2\, ds\right).$$

5 En prenant $\varphi = 1_{\|\cdot\|_{L^\infty(0,t)} \leq \varepsilon}$ dans (2), en déduire que

$$\lim_{\varepsilon \to 0} \frac{\mathbb{P}(\|W_s - h(s)\|_{L^\infty(0,t)} \leq \varepsilon)}{\mathbb{P}(\|W_s\|_{L^\infty(0,t)} \leq \varepsilon)} = \exp\left(-\frac{1}{2}\int_0^t (h'(s))^2\, ds\right).$$