Université Paris Dauphine
Département MIDO
2020–2021
Master 1 – Méthodes de Monte Carlo

Examen Final – Janvier 2021

DURÉE 3H00 – DOCUMENTS ET CALCULATRICE NON AUTORISÉS

Important. Suivant les réglements en vigueur :

1. Les enseignants présents lors de l'épreuve ne peuvent communiquer que sur les fautes d'énoncé potentielles. Toute autre question durant la composition ne sera pas acceptée.
2. Les étudiants sont tenus de se lever au moment de l'annonce de fin de la composition. En cas de refus, le responsable de l'UE sera fondé à ne pas prendre en compte la copie incriminée.
3. L'identification des copies et intercalaires doit avoir été faite au moment de la remise de chaque copie par les enseignants et surveillants. Il ne sera pas accordé de délai pour cette raison en fin d'épreuve.

Notation. *i.i.d.* : indépendantes et identiquement distribuées.

Exercice 1. Soit $\mathbf{X}$ un vecteur aléatoire de $\mathbb{R}^d$ de densité f. Soient $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ et $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ mesurables telles que $\phi(\mathbf{X})$ et $\psi(\mathbf{X})$ sont de carré intégrables. On souhaite estimer

$$\Delta = \mathbb{E}\big[\phi(\mathbf{X}) - \psi(\mathbf{X})\big].$$

Étant données $\mathbf{X}_1, \dots, \mathbf{X}_{2n}$, $n \in \mathbb{N}^*$, variables aléatoires *i.i.d.* suivant f, on considère les estimateurs

$$\widehat{D}_n = \frac{1}{n}\sum_{k=1}^{n}\big[\phi(\mathbf{X}_k) - \psi(\mathbf{X}_k)\big] \quad \text{et} \quad \widehat{\Delta}_n = \frac{1}{n}\sum_{k=1}^{n}\phi(\mathbf{X}_k) - \frac{1}{n}\sum_{k=1}^{n}\psi(\mathbf{X}_{n+k}).$$

On note C le coût de simulation d'une réalisation de f et on suppose que le coût de calcul de $\widehat{D}_n$ et $\widehat{\Delta}_n$ est déterminé par le nombre de réalisations simulées suivant f.

1. Montrer que $\widehat{D}_n$ et $\widehat{\Delta}_n$ sont des estimateurs sans biais de Δ.

2. Quel est le coût de calcul nécessaire pour obtenir une erreur quadratique moyenne $\varepsilon \in \mathbb{R}_+$ avec $\widehat{D}_n$? avec $\widehat{\Delta}_n$?

3. À quelle condition nécessaire et suffisante sur $\mathrm{Cov}\big[\phi(\mathbf{X}), \psi(\mathbf{X})\big]$, $\widehat{D}_n$ atteint-il la même erreur quadratique moyenne que $\widehat{\Delta}_n$ pour un coût moindre ?

Application. Soient des variables aléatoires $Y_1, \dots, Y_d$ *i.i.d.* suivant la loi $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$, et $Z_1, \dots, Z_d$ *i.i.d.* suivant la loi $\mathcal{N}(\lambda, \gamma^2)$, $\lambda \in \mathbb{R}$, $\gamma \in \mathbb{R}_+^*$. Pour h mesurable et bornée sur $\mathbb{R}$, on s'intéresse à l'estimation de

$$\eta = \mathbb{E}\left[\sum_{k=1}^{d} h(Y_k) - \sum_{k=1}^{d} h(Z_k)\right].$$

4. Montrer que pour estimer η, il est préférable, en terme de variance et de coût, d'utiliser comme réalisations de $Z_1, \dots, Z_d$ une transformation des simulations de $Y_1, \dots, Y_d$ dès que $\mathrm{Cov}[h(Y_1), h(Z_1)] > 0$,

plutôt que de simuler $Y_1,\dots,Y_d$ et $Z_1,\dots,Z_d$ indépendamment.

Exercice 2. Soit $n\in\mathbb{N}$ tel que $n\geq 2$. Soient $X_1,\dots,X_n$ des variables aléatoires réelles *i.i.d.* de densité f et de fonction de répartition F dont l'inverse généralisée $F^{\leftarrow}$ est facilement calculable. Les statistiques d'ordre $X_{(1)},\dots,X_{(n)}$ sont obtenues en faisant un tri croissant de $X_1,\dots,X_n$. Pour $k\in\{1,\dots,n\}$, la densité de $X_{(k)}$ est donnée, pour $x\in\mathbb{R}$, par

$$f_k(x)=n!\frac{F(x)^{k-1}}{(k-1)!}\frac{[1-F(x)]^{n-k}}{(n-k)!}f(x),$$

et celle du couple $(X_{(1)},X_{(n)})$ est donnée, pour $x,y\in\mathbb{R}$ tels que $x\leq y$ par

$$f_{1,n}(x,y)=n(n-1)[F(y)-F(x)]^{n-2}f(x)f(y).$$

1. **(a)** Soit $1<k<n$. Montrer que pour simuler une réalisation de $X_{(k)}$, l'algorithme du rejet avec f pour densité instrumentale permet de simuler en moyenne moins de réalisations suivant f qu'une approche directe basée sur le tri de n réalisations suivant f.

Indication : on pourra utiliser sans le démontrer que pour $1<k<n$,

$$(n-1)!(k-1)^{k-1}(n-k)^{n-k}\leq\frac{1}{2}(n-1)^{n-1}(k-1)!(n-k)!$$

(b) Dans R, on a enregistré une M tel que, pour $x\in\mathbb{R}$, $f_k(x)\leq Mf(x)$ et `rho(x)` une fonction qui calcule (de façon vectorielle) le rapport $f_k(x)/[Mf(x)]$ pour $X_1,\dots,X_n$ distribuées (dans cette question uniquement) suivant la loi exponentielle $\mathscr{E}(2)$. Donner le code R de l'algorithme du rejet qui simule N réalisations de $X_{(k)}$. Ce code doit être optimisé pour minimiser le nombre de passages dans une boucle `while`. On pourra utiliser les fonctions de la librairie de base de R.

2. **(a)** Pour $y\in\mathbb{R}$, déterminer la fonction de répartition de $X_{(1)}\mid X_{(n)}=y$, en fonction de F.

(b) En déduire une méthode de simulation du couple $(X_{(1)},X_{(n)})$ basée sur F, $F^{\leftarrow}$ et sur la simulation de moins de n variables uniformes sur $[0,1]$.

(c) Donner, pour cette méthode, le code R qui fournit une réalisation `(x,y)` du couple $(X_{(1)},X_{(n)})$ lorsque $X_1,\dots,X_n$ sont distribuées suivant la loi de Cauchy $\mathscr{C}(0,1)$.

Exercice 3. Soit $(X_k)_{k\geq 1}$ une suite de variables aléatoires réelles *i.i.d.* suivant une densité f paramétrée par $\theta\in\Theta\subseteq\mathbb{R}$, où Θ est un intervalle contenant 0. On suppose que f vérifie l'hypothèse (H) : pour $\theta\in\Theta$ et $x\in\mathbb{R}$,

$$f(x\mid\theta)=h(x)\exp(\theta x-\psi(\theta))\quad\text{avec}\quad\begin{cases}h:\mathbb{R}\mapsto\mathbb{R}_+^*\\ \psi:\Theta\mapsto\mathbb{R}\end{cases}\qquad\text{telle que}\begin{cases}\psi\in\mathscr{C}^2(\Theta),\text{ avec }\psi(0)=\psi'(0)=0\\ \forall\theta\in\Theta,\ \psi'(\theta)=\mathbb{E}_\theta[X_1],\\ \forall\theta\in\Theta,\ \psi''(\theta)=\mathbb{V}\mathrm{ar}_\theta[X_1].\end{cases}$$

Soient $a, b \in \mathbb{R}$ tels que $a \leq 0 < b$. On définit

$$S_m = \begin{cases} 0 & \text{si } m = 0 \\ \sum_{k=1}^{m} X_k & \text{si } m \geq 1 \end{cases} \quad \text{et} \quad M = \min\left\{ m \in \mathbb{N}^* \;\middle|\; S_m \notin]a, b[\right\}.$$

Soit $\theta_0 \in \Theta$ tel que $\mathbb{E}_{\theta_0}[X_1] < 0$ et $\mathbb{P}_{\theta_0}[X_1 > 0] > 0$, où $\mathbb{P}_{\theta_0}$ et $\mathbb{E}_{\theta_0}$ désignent respectivement la probabilité et l'espérance par rapport à $f(\cdot \mid \theta_0)$. On souhaite estimer $p = \mathbb{P}_{\theta_0}[S_M \geq b]$. On notera $n \in \mathbb{N}^*$ le nombre de tirages suivant le modèle.

1. Donner l'expression de l'estimateur de Monte Carlo classique de p, noté $\widehat{p}_n(\theta_0)$, pour des variables aléatoires (X_{ij}), $i \in \{1, \ldots, n\}$ et $j \in \mathbb{N}^*$, *i.i.d.* suivant $f(\cdot \mid \theta_0)$.

2. Soit $\theta \in \Theta \setminus \{\theta_0\}$. Donner l'estimateur d'échantillonnage préférentiel de p pour des variables aléatoires (X_{ij}), $i \in \{1, \ldots, n\}$ et $j \in \mathbb{N}^*$, *i.i.d.* suivant $g : x \mapsto f(x \mid \theta)$, noté $\widehat{\delta}_n(\theta)$.

3. Montrer que l'on a nécessairement $\theta_0 < 0$ et qu'il existe au plus un point $\theta^\star > 0$ tel que $\psi(\theta^\star) = \psi(\theta_0)$.

4. Montrer que $\mathbb{V}\mathrm{ar}\left[\widehat{\delta}_n(\theta^\star)\right] \leq \mathbb{V}\mathrm{ar}\left[\widehat{p}_n(\theta_0)\right]$.

Application. On suppose que $(X_n)_{n \geq 1}$ sont *i.i.d.* suivant la loi $\mathcal{N}(\theta_0, 1)$ avec $\theta_0 < 0$. On considère $\Theta = \mathbb{R}$.

5. Montrer que la densité de la loi $\mathcal{N}(\theta_0, 1)$ vérifie l'hypothèse (H) et donner l'expression de $\theta^\star$.

6. En déduire un majorant de p en fonction de θ_0 et b.

7. On prend $a = -5$, $b = 10$ et $\theta_0 = -1$. Pour $\widehat{p}_n(\theta_0)$, à partir de quelle valeur de n peut-on espérer avoir une estimation de l'ordre de p ?

x	-20	-15	-10	-5
$\exp(x)$	2.061×10^{-9}	3.059×10^{-7}	4.54×10^{-5}	0.007

Exercice 4. Soit $d \in \mathbb{N}^*$. Soient $X_1, \ldots, X_d$ des variables aléatoires *i.i.d.* suivant la loi arc sinus standard dont la densité est donnée, pour $x \in \mathbb{R}$, par

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{\{x \in [0,1]\}}.$$

On s'intéresse à l'estimation de

$$\delta = \mathbb{E}[h(X_1, \ldots, X_d)], \quad \text{avec} \quad h : (x_1, \ldots, x_d) \mapsto \log\left(\frac{1}{d} \sum_{i=1}^{d} e^{X_i} \right).$$

1. Calculer l'inverse généralisé de la fonction de répartition de la loi arc sinus standard et expliquer comment obtenir des réalisations de cette loi.

2. Soit $n \in \mathbb{N}^*$. Donner l'expression de l'estimateur de Monte Carlo classique de δ, noté $\widehat{\mu}_n$, pour des variables aléatoires (X_{ij}), $i \in \{1, \ldots, d\}$ et $j \in \{1, \ldots, n\}$, *i.i.d.* suivant f.

3. **(a)** Trouver des réels a et b tels que $aX_1 + b$ a même loi que X_1.

(b) En déduire l'estimateur de δ, noté $\widehat{\delta}_n$, par la méthode de la variable antithétique.

(c) Montrer que $\mathbb{V}\mathrm{ar}\big[\widehat{\delta}_n\big] \leq \mathbb{V}\mathrm{ar}\big[\widehat{\mu}_n\big]/2$.

4. On pose $h_0 : (x_1,\dots,x_d) \mapsto \sum_{i=1}^d x_i$. Montrer que $\mathbb{E}[h_0(X_1,\dots,X_d)] = d/2$. En déduire, en fonction d'un paramètre $b \in \mathbb{R}$ et de h_0, un estimateur par la méthode de la variable de contrôle simple, noté $\widehat{\delta}_n(b)$.

5. Numériquement, on observe que $\widehat{\delta}_n$ a une variance 60 fois plus faible que $\widehat{\mu}_n$ et que le coefficient de corrélation au carré entre $h_0(X_1,\dots,X_d)$ et $h(X_1,\dots,X_d)$ vaut $\rho^2 = 0.98$. Soit $b^\star \in \mathbb{R}$ tel que $\widehat{\delta}_n(b^\star)$ soit de variance minimale. Comparer la variance de $\widehat{\delta}_n$ et celle de $\widehat{\delta}_n(b^\star)$.

Exercice 5. Soit N une variable aléatoire de loi de Poisson $\mathscr{P}(\lambda)$, $\lambda \in \mathbb{R}_+^*$, *i.e.*, pour $k \in \mathbb{N}$, $\mathbb{P}[N=k] = \lambda^k e^{-\lambda}/(k!)$. Soient $Z = N \mid N \geq 1$ et $(W_t)_{t\in\mathbb{R}_+}$ un mouvement brownien standard. On s'intéresse à

$$\Delta = \mathbb{E}[\max(W_0, W_1, \dots, W_Z)].$$

1. Donner le code R, sans boucles ou fonctions `apply`, qui fournit une réalisation de $(W_0, W_1, \dots, W_Z)$, pour $\lambda = 2$.

2. Soient $n \in \mathbb{N}^*$ et $L \in \mathbb{N}$ tel que $L \geq 2$. On pose pour $k \in \{1,\dots,L-1\}$

$$n_k = \frac{n\lambda^k}{k!(e^\lambda - 1)} \quad \text{et} \quad n_L = n - \sum_{k=1}^{L-1} n_k.$$

Pour des variables aléatoires $X_{i,\ell}^{(k)}$, $k \in \{1,\dots,L\}$, $i \in \{1,\dots,n_k\}$, $\ell \in \mathbb{N}^*$, *i.i.d.* suivant la loi $\mathscr{N}(0,1)$ et $Y_1,\dots,Y_{n_L}$ *i.i.d.* suivant la loi de $Z \mid Z \geq L$, on définit

$$W_{i,j}^{(k)} = \sum_{\ell=1}^{j} X_{i,\ell}^{(k)}, \quad \text{avec} \quad k \in \{1,\dots,L\},\ i \in \{1,\dots,n_k\},\ j \in \mathbb{N}^*, \quad \text{et}$$
$$\widehat{p}_n(n_1,\dots,n_L) = \frac{1}{n}\sum_{k=1}^{L-1}\sum_{i=1}^{n_k} \max\Big(0, W_{i,1}^{(k)}, \dots, W_{i,k}^{(k)}\Big) + \frac{1}{n}\sum_{i=1}^{n_L} \max\Big(0, W_{i,1}^{(L)}, \dots, W_{i,Y_i}^{(L)}\Big).$$

Montrer que $\widehat{p}_n(n_1,\dots,n_L)$ est un estimateur stratifié avec allocation proportionnelle pour n tirages au total et L strates, notées $D_1,\dots,D_L$.

3. *(Question de cours)* Pour la variable de stratification Z et les strates $D_1,\dots,D_L$, montrer que l'estimateur stratifié de variance minimale est obtenu pour l'allocation $(n_1,\dots,n_L)$ définie par

$$n_k = n\frac{\mathbb{P}[Z \in D_k]\sqrt{\mathbb{V}\mathrm{ar}[\max(W_0, W_1, \dots, W_Z) \mid Z \in D_k]}}{\sum_{i=1}^L \mathbb{P}[Z \in D_i]\sqrt{\mathbb{V}\mathrm{ar}[\max(W_0, W_1, \dots, W_Z) \mid Z \in D_i]}}, \quad k \in \{1,\dots,L\},$$

et que sa variance est inférieure à $\mathbb{V}\mathrm{ar}[\max(W_0, W_1, \dots, W_Z)]/n$.