Dauphine | PSL
UNIVERSITÉ PARIS

MIDO - Master 1 Mathématiques & Applications, 2022–2023

Méthodes de Monte Carlo

Examen Final - 11/01/2023

DURÉE 3H00 – DOCUMENTS ET CALCULATRICE NON AUTORISÉS

Ex. 1	Ex. 2	Ex. 3	Ex. 4	Total
...... / 5.5	...... / 4.5	...... / 5	...... / 5	...... / 20

NOM : ....

PRÉNOM : ....

(lisiblement)

PARTIE À RABATTRE

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 de la copie de composition doit se faire au moment de la remise de la copie par les enseignants et surveillants. Il ne sera pas accordé de délai pour cette raison en fin d'épreuve.

Les exercices sont indépendants. Toutes les réponses sont à fournir sur la copie d'énoncé. L'espace blanc alloué à chaque question est suffisant pour apporter une réponse correcte. Un espace complémentaire est disponible en fin de copie.

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

Exercice 1

...... / **5.5**

Soient f et g des densités par rapport à la mesure de Lebesgue sur $\mathbb{R}$ de la forme

$$\forall x \in \mathbb{R}, \quad f(x) = c\tilde{f}(x) \quad \text{et} \quad g(x) = d\tilde{g}(x),$$

où les fonctions $\tilde{f}$ et $\tilde{g}$ sont positives, connues et faciles à calculer, mais où les constantes positives c et d sont inconnues. **Pour les questions 1. à 4.**, on considère le cas particulier d'un intervalle $]a,b[\subset \mathbb{R}$ tel que

$$\forall x \notin]a,b[, \quad \tilde{f}(x) = 0 \quad \text{et} \quad \sup_{x\in\mathbb{R}} \tilde{f}(x) = M < \infty.$$

1. Pour un point aléatoire $U = (U_1, U_2)$ tiré uniformément dans $\mathscr{R} =]a,b[\times]0,M[$, calculer la probabilité $\mathbb{P}\left[U_2 \leq \tilde{f}(U_1)\right]$.

...... / **0.5**

PARTIE À RABATTRE

2. Pour une suite $U^1, \dots, U^n$, $n \in \mathbb{N}^*$, de variables aléatoires *i.i.d.* de loi uniforme sur $\mathscr{R}$, construire un estimateur convergent de c. Vous justifierez soigneusement la convergence de l'estimateur.

...... /**1**

3. L'estimateur ainsi construit est-il sans biais ?

...... /**0.5**

4. Avec R, on dispose de la fonction vectorisée `df` permettant d'évaluer la fonction $\tilde{f}$. Écrire un code R qui calcule cette estimation de c.

...... /**1**

À partir de maintenant, il n'est plus nécessaire de supposer que f est bornée et admet un support borné. On suppose néanmoins jusqu'à la fin de l'exercice que $\{x \in \mathbb{R} \mid f(x) = 0\} \subseteq \{x \in \mathbb{R} \mid g(x) = 0\}$.

5. Soit X une variable aléatoire de densité f. Montrer que

$$\mathbb{E}\left[\frac{\tilde{g}(X)}{\tilde{f}(X)}\right] = \frac{c}{d}.$$

...... /**0.5**

6. Pour une suite $X_1, \ldots, X_n$ de variables aléatoires *i.i.d.* de densité f, déduire de la question précédente un estimateur convergent de c/d.

...... /**0.5**

7. Soit $\alpha(\cdot)$ une fonction strictement positive sur $\mathbb{R}$ telle que

$$\int_{\mathbb{R}} \alpha(x)\tilde{f}(x)\tilde{g}(x)\mathrm{d}x < +\infty.$$

Montrer que si X est une variable aléatoire de densité f et Y une variable aléatoire de densité g, alors

$$\frac{\mathbb{E}\left[\alpha(X)\tilde{g}(X)\right]}{\mathbb{E}\left[\alpha(Y)\tilde{f}(Y)\right]} = \frac{c}{d}.$$

...... /**1**

8. Deduire de la question précédente un estimateur convergent de c/d basé sur deux suites $X_1, \dots, X_n$ et $Y_1, \dots, Y_n$ de variables aléatoires *i.i.d.* de densité f et g respectivement. Vous justifierez la convergence de l'estimateur.

...... $/$**0.5**

Exercice 2

...... $/$**4.5**

Soient $m \in \mathbb{N}^*$ et $k \in [\![2, m-1]\!]$. Soit une densité f. On note F la fonction de répartition associée et $F^{\leftarrow}$ l'inverse généralisé de F. On considère la variable aléatoire X de densité

$$g(x) = \frac{m!}{(k-1)!(m-k)!} F(x)^{k-1} [1 - F(x)]^{m-k} f(x).$$

On souhaite simuler suivant g en utilisant l'algorithme du rejet avec f comme densité instrumentale.

1. Montrer qu'en moyenne, cet algorithme (bien optimisé) nécessite moins de $m/2$ tirages suivant f pour obtenir un tirage suivant g.

Indication. on pourra utiliser sans le démontrer que pour $1 < k < m$,

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

...... $/$**1**

On considère (U, V) un couple de variables aléatoires dont la densité jointe est donnée, pour $u, v \in \mathbb{R}$ par

$$h(u, v) = n(n-1)[F(v) - F(u)]^{n-2} f(u) f(v) \mathbb{1}_{\{u \leq v\}}.$$

2. Montrer que la densité marginale de V est

$$\phi(v) = nF(v)^{n-1} f(v).$$

...... /**0.5**

3. En déduire la fonction de répartition de V et une méthode de simulation de V à partir du générateur aléatoire de la loi uniforme sur $[0,1]$.

...... /**0.5**

4. Pour $v \in \mathbb{R}$, déterminer la fonction de répartition de $U \mid V = v$ en fonction de F.

...... /**1**

5. En déduire une méthode de simulation du couple (U, V) à partir du générateur aléatoire de la loi uniforme sur $[0,1]$.

...... /**1**

6. Donner le code R qui fournit une réalisation `(u,v)` du couple (U, V) lorsque f est la densité de la loi de Poisson de paramètre $\lambda = 4$.

...... /**0.5**

Exercice 3

...... /**5**

On souhaite estimer à l'aide de méthodes de Monte Carlo l'aire délimitée par une cardioïde :

$$\delta = \frac{1}{2}\int_0^{2\pi} [1 - \cos(\theta)]^2 \, \mathrm{d}\theta.$$

1. Écrire un estimateur Monte Carlo $\widehat{\delta}_n$ de δ utilisant une suite de $n \in \mathbb{N}^*$ variables aléatoires $(U_k)_{k\geq 1}$ *i.i.d.* suivant la loi uniforme sur $[0,2\pi]$.

...... /**0.5**

2. Donner le code R d'une fonction `mc(n)` qui fournit, pour `n` tirages suivant la loi uniforme sur $[0,2\pi]$, l'estimation $\widehat{\delta}_n$, la variance $\mathbb{V}\mathrm{ar}\left[\widehat{\delta}_n\right]$, ainsi qu'un intervalle de confiance au niveau 90%. La sortie de la fonction doit être sous forme d'une liste avec les arguments `value` pour l'estimation, `var` pour la variance et `ic` pour l'intervalle de confiance.

...... /**1**

3. Montrer que si U suit la loi uniforme sur $[0,2\pi]$, alors pour tout $k \in [\![1,n]\!]$,

$$\frac{2\pi(k-1)+U}{n} \quad \text{a la même loi que } U \text{ sachant que} \quad U \in \left]\frac{2\pi}{n}(k-1), \frac{2\pi}{n}k\right].$$

...... /**1**

4. Montrer que si U suit la loi uniforme sur $[0,2\pi]$, alors

$$\mathbb{V}\mathrm{ar}\left[[1-\cos(U)]^2\right] \geq \frac{1}{n}\sum_{k=1}^{n} \mathbb{V}\mathrm{ar}\left[\left[1-\cos\left(\frac{2\pi(k-1)+U}{n}\right)\right]^2\right]$$

...... /1

5. En déduire un estimateur $\widehat{\Delta}_n$ de δ utilisant une suite de $n \in \mathbb{N}^*$ variables aléatoires $(U_k)_{k\geq 1}$ *i.i.d.* suivant la loi uniforme sur $[0,2\pi]$ et tel que $\mathbb{V}\mathrm{ar}\left[\widehat{\Delta}_n\right] \leq \mathbb{V}\mathrm{ar}\left[\widehat{\delta}_n\right]$.

...... / **0.5**

6. Donner le code R d'une fonction `mc(n)` qui fournit, pour `n` tirages suivant la loi uniforme sur $[0,2\pi]$, l'estimation $\widehat{\Delta}_n$ et la variance $\mathbb{V}\mathrm{ar}\left[\widehat{\Delta}_n\right]$. La sortie de la fonction doit être sous forme d'une liste avec les arguments `value` pour l'estimation et `var` pour la variance.

...... / **1**

Exercice 4

...... /**5**

Soit $\mu \in \mathbb{R}$. On considère X la variable aléatoire de densité, par rapport à la mesure de Lesbesgue sur $\mathbb{R}$,

$$f(x) = \frac{3}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-\mu)^2\right\} - \frac{2}{\sqrt{\pi}} \exp\left\{-(x-\mu)^2\right\}.$$

On suppose que l'on sait simuler une suite $X_1, \ldots X_n$ de variable aléatoire *i.i.d.* suivant f à l'aide de l'algorithme du rejet. Pour une fonction h continue, monotone et bornée sur $\mathbb{R}$, on souhaite estimer

$$\delta = \int_{\mathbb{R}} x h(x) f(x) \mathrm{d}x.$$

1. Montrer que $Xh(X)$ est de carré intégrable, *i.e.*, $\mathbb{E}\left[X^2 h(X)^2\right] < \infty$.

...... /**0.5**

2. Montrer qu'il existe une transformation affine $A : \mathbb{R} \mapsto \mathbb{R}$ différente de l'identité telle que $A(X)$ et X ont même loi.

...... /**1**

3. En déduire un estimateur $\widehat{\delta}_n$ de δ par la méthode de la variable antithétique.

...... / **0.5**

4. Montrer que $n\mathbb{V}\text{ar}\left[\widehat{\delta}_n\right] \leq \mathbb{V}\text{ar}[Xh(X)]$.

...... / **1**

5. Pour tout $k \in \mathbb{N}^*$, calculer $\mathbb{E}\left[(X-\mu)^k\right]$.

Indication. On rappelle que pour une variable aléatoire gaussienne Z de moyenne μ et de variance σ^2 on a

$$\mathbb{E}\left[(Z-\mu)^{2k+1}\right] = 0 \quad \text{et} \quad \mathbb{E}\left[(Z-\mu)^{2k}\right] = \frac{(2k)!}{k!}\left(\frac{\sigma^2}{2}\right)^k.$$

...... / **0.5**

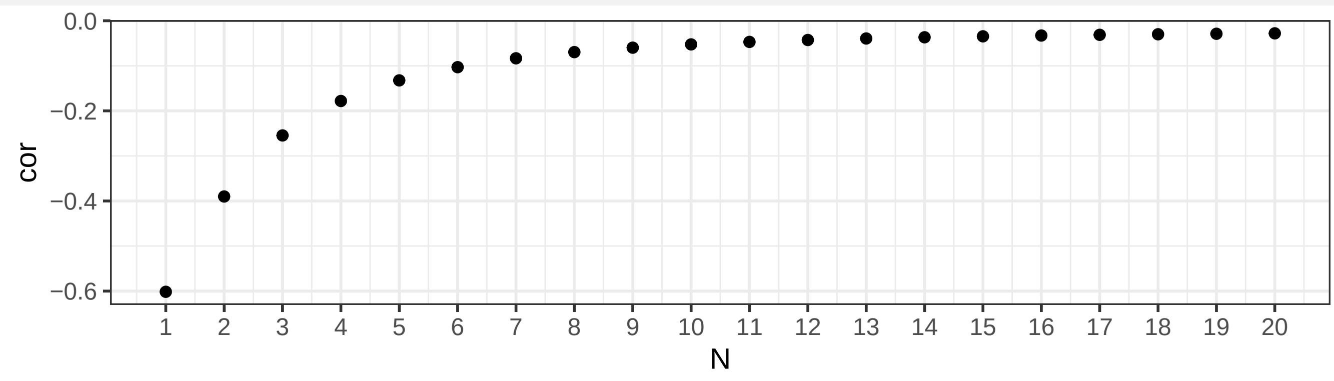
Soit $N \in \mathbb{N}^*$. On considère la fonction

$$\begin{array}{rcccl} h_0 & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & \sum_{i=1}^{2N} (x-\mu)^i \end{array}$$

6. Donner un estimateur $\widehat{\Delta}_n$ de δ par la méthode de la variable de contrôle.

...... / **0.5**

7. Le graphe ci-dessous représente la corrélation entre $Xh(X)$ et $h_0(X)$ en fonction de N. Quelle valeur de N choisiriez-vous pour votre estimateur par la méthode de la variable de contrôle? Justifier votre choix.



...... / **0.5**

8. Soit le vecteur

$$\texttt{v} = (\texttt{var\_mc, var\_anti, var\_cont}) = \left(\frac{1}{n}\mathbb{V}\mathrm{ar}[Xh(X)], \mathbb{V}\mathrm{ar}\left[\widehat{\delta}_n\right], \mathbb{V}\mathrm{ar}\left[\widehat{\Delta}_n\right]\right).$$

La commande R : `outer(v, v, '/')` donne

	var_mc	var_anti	var_cont
var_mc	1.000000	23.603	1.567085
var_anti	0.042367	1.000	0.066393
var_cont	0.638127	15.062	1.000000

Quelle est la méthode la plus efficace en terme de variance?

...... / **0.5**

Espace complémentaire