

Méthodes de Monte Carlo – Projet

stoehr@ceremade.dauphine.fr

- **À rendre avant le 23 décembre 2021. Chaque jour de retard sera pénalisé d'un point.**
- **R est le seul langage autorisé.** Les questions nécessitant un code R sont indiquées par le symbole ♠.
- Le rapport (**nom du fichier : numero_groupe_rapport_noms**, e.g., 001_rapport_Stoehr_Maillet)
 - à rendre au format **.pdf** et doit contenir vos réponses et commentaires. Une rédaction soignée est attendue. Il est important de justifier/commenter les résultats théoriques et numériques
 - Pour intégrer tout ou partie de votre code et des sorties dans votre rapport, vous pouvez utiliser les outils dédiés : Notebook, Rmarkdown ou \LaTeX + knitr. En revanche, **il est interdit de copier-coller du code brut dans le corps du texte.**
 - Les graphiques doivent être soigneusement annotés et présentés (titre, couleur, légendes, ...).
- Une version du code pouvant être testée doit être fournie (**même nom de fichier**). Ce code doit
 - s'exécuter sans erreurs et permettre de reproduire l'intégralité des résultats présentés dans le rapport. Vous préciserez la graine utilisée pour les résultats obtenus.
 - être bien commenté. Il est possible qu'une explication orale vous soit demandée.
 - utiliser autant que possible les spécificités du langage (bonus pour les codes les plus efficaces).

Exercice 1 (*Déambulation vers la sortie*). On considère la marche aléatoire définie pour $m \in \mathbb{N}$ par

$$W_m = \begin{cases} 0 & \text{si } m = 0 \\ \sum_{k=1}^m X_k & \text{si } m \geq 1 \end{cases}$$

avec $(X_n)_{n \in \mathbb{N}}$ une suite de variables *i.i.d.* de loi $\mathcal{N}(\mu, 1)$, $\mu \in \mathbb{R}^*$. Cette marche aléatoire sera observée pendant une durée aléatoire

$$T = \inf\{m \in \mathbb{N}^* \mid W_m \notin]a, b[\}, \quad \text{avec } a, b \in \mathbb{R} \text{ tels que } a \leq 0 < b.$$

On souhaite estimer $\delta = \mathbb{P}_\mu[W_T \geq b]$. On notera $n \in \mathbb{N}^*$ le nombre de tirages de variables aléatoires effectués.

1. On définit $S = \inf\{m \in \mathbb{N}^* \mid |X_m| > b - a\}$.

(a) Montrer que $\mathbb{P}[S > n] \xrightarrow{n \rightarrow +\infty} 0$.

(b) En déduire que T est presque sûrement fini.

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

3. (a) Soit $\theta \in \mathbb{R} \setminus \{\mu\}$. Donner l'estimateur d'échantillonnage préférentiel de δ pour des variables aléatoires X_{ij} , $i \in \llbracket 1, n \rrbracket$ et $j \in \mathbb{N}^*$, *i.i.d.* de loi $\mathcal{N}(\theta, 1)$, noté $\widehat{\Delta}_n(\theta)$.

(b) Montrer que pour $\theta = -\mu$, on a $\text{Var}[\widehat{\Delta}_n(-\mu)] \leq \text{Var}[\widehat{\delta}_n]$.

Application. On suppose que $a = -3$, $b = 5$, $\mu = -1$ et $n = 10^6$.

4. Justifier le choix de n à l'aide d'une majoration de δ en fonction de μ et b .
5. (♠) Calculer $\widehat{\delta}_n$ et $\widehat{\Delta}_n(-\mu)$. Déterminer le rapport des variances et l'efficacité relative de ces deux méthodes. Discuter le résultat obtenu.

Exercice 2 (Recettes de financier).

Motivation. Le modèle de Black–Scholes est utilisé pour modéliser la dynamique d'un marché financier, le prix de l'action étant représenté par un processus stochastique en temps continu. Le modèle s'écrit $W_t(B) = W_0 \exp[(r - 0.5\sigma^2)t - \sigma B_t]$, où $W_0 \in \mathbb{R}_+^*$ est la valeur actuelle de l'option sous-jacente, $r \in \mathbb{R}_+^*$ le taux d'intérêt, $\sigma \in \mathbb{R}_+^*$ la volatilité du prix de l'action et $(B_t)_{t \in \mathbb{R}_+}$ un mouvement brownien standard. À l'aide de ce modèle, on peut par exemple définir le prix théorique d'une option d'achat : $\mathbb{E}[\exp(-rT) \max(W_T(B) - K, 0)]$ où $T \in \mathbb{R}_+^*$ est la durée avant échéance de l'option et $K \in \mathbb{R}_+^*$ est le prix d'exercice (*strike*) de l'option. L'estimation du prix de l'option d'achat peut donc se faire simplement via les méthodes de Monte Carlo et un générateur suivant la loi normale.

Dans cet exercice, on ne se limitera pas au seul problème de l'estimation du prix d'une option d'achat mais les différentes méthodes étudiées sont utilisées dans ce domaine d'application.

Partie I. Soit $(W_n)_{n \in \mathbb{N}}$ la marche aléatoire introduite dans l'Exercice 1. On souhaite estimer pour $T \in \mathbb{N}^*$ fixé

$$\delta = \mathbb{E}[\max\{\lambda \exp(-\sigma W_T) - K, 0\}], \quad \text{avec } \lambda, \sigma, K \in \mathbb{R}_+^*.$$

Pour les applications numériques, on prendra $\mu = -1$, $T = 91$, $\lambda = 30.63$, $\sigma = 0.01$, $K = 72.36$ et $n = 10^5$ tirages Monte Carlo.

1. (a) Donner la loi ν de W_T . En déduire l'expression de l'estimateur de Monte Carlo classique de δ , noté $\widehat{\delta}_n$, pour des variables aléatoires Y_i , $i \in \llbracket 1, n \rrbracket$ suivant la loi ν .
 (b) (♠) Calculer $\widehat{\delta}_n$ ainsi que l'erreur quadratique moyenne relativement à δ .
2. (a) En utilisant une transformation A qui laisse la loi ν invariante, donner l'expression de l'estimateur de δ par la méthode de la variable antithétique, noté $\widehat{\delta}_n(A)$ et montrer que $\text{Var}[\widehat{\delta}_n] \geq \text{Var}[\widehat{\delta}_n(A)]$.
 (b) (♠) Calculer $\widehat{\delta}_n(A)$ ainsi que l'erreur quadratique moyenne relativement à δ .
3. (♠) Déterminer l'efficacité relative de ces deux méthodes d'estimation de δ . Commenter le résultat.

Partie II. Soit $\mathbf{X} = (X_1, X_2)$ un vecteur gaussien centré tel que $\text{Var}[X_1] = \text{Var}[X_2] = 1$ et $\text{Cov}[X_1, X_2] = \rho \in]-1, 1[$. On souhaite estimer

$$\mathcal{I} = \mathbb{E}[\max\{\lambda_1 \exp(-\sigma_1 X_1) + \lambda_2 \exp(-\sigma_2 X_2) - K, 0\}], \quad \text{avec } \lambda_1, \lambda_2, \sigma_1, \sigma_2, K \in \mathbb{R}_+^*.$$

Pour les applications numériques, on prendra $\rho = -0.72$, $(\lambda_1, \lambda_2) = (32, 35.9)$, $(\sigma_1, \sigma_2) = (0.5, 0.3)$, $K = 67.8$ et $n = 10^4$ tirages Monte Carlo.

4. (a) Donner l'expression de l'estimateur de Monte Carlo classique de \mathcal{I} , noté $\widehat{\mathcal{I}}_n$, pour des variables aléatoires $\mathbf{X}_i = (X_{1i}, X_{2i})$, $i \in \llbracket 1, n \rrbracket$ suivant la loi de \mathbf{X} .

- (b) (♠) Calculer $\hat{\mathcal{I}}_n$ ainsi qu'un intervalle de confiance asymptotique bilatère au niveau 90%.
- 5. (a) Calculer explicitement $\mathbb{E}[\exp(-\sigma_1 X_1)]$, $\mathbb{E}[\exp(-\sigma_2 X_2)]$ et $\mathbb{E}[\lambda_1 \exp(-\sigma_1 X_1) + \lambda_2 \exp(-\sigma_2 X_2)]$.
- (b) (♠) Proposer une estimation de \mathcal{I} par la méthode de la variable de contrôle, notée $\hat{\mathcal{I}}_n(b)$, $b \in \mathbb{R}$, avec un intervalle de confiance asymptotique bilatère au niveau 90%. Justifier avec soin vos différents choix.
- 6. (♠) Déterminer l'efficacité relative de ces deux méthodes d'estimation de \mathcal{I} . Commenter le résultat.

Partie III. Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur gaussien standard de \mathbb{R}^d . On souhaite estimer

$$p = \mathbb{E} \left[\max \left\{ \frac{1}{d} \sum_{i=1}^d \exp(\sigma X_i) - K, 0 \right\} \right], \quad \text{avec } \sigma, K \in \mathbb{R}_+^*.$$

Pour les applications numériques, on prendra $d = 10$, $\sigma = 0.5$, $K = 1.17$ et $n = 10^5$ tirages Monte Carlo.

- 7. (♠) Donner une estimation de p , notée \hat{p}_n par la méthode de Monte Carlo classique, *i.e.*, basée sur le générateur de la loi normale $\mathcal{N}(0, 1)$, et de la variance de la méthode.

Dans la suite, note $\langle \cdot, \cdot \rangle$ le produit scalaire usuel de \mathbb{R}^d et $\| \cdot \|$ la norme euclidienne associée.

- 8. Soient $u \in \mathbb{R}^d$ tel que $\|u\|^2 = 1$. Montrer que \mathbf{X} a même loi que $\langle \mathbf{X}, u \rangle u + \mathbf{Z} - \langle \mathbf{Z}, u \rangle u$, avec $\mathbf{Z} = (Z_1, \dots, Z_d)^\top$ un vecteur gaussien standard de \mathbb{R}^d indépendant de \mathbf{X} .
- 9. Soit $L \in \mathbb{N}^*$ fixé. On considère (D_1, \dots, D_L) une partition de \mathbb{R} . Pour $\ell \in \llbracket 1, L \rrbracket$, on note Y_ℓ la variable aléatoire qui suit la loi de $\langle \mathbf{X}, u \rangle$ sachant que $\langle \mathbf{X}, u \rangle \in D_\ell$. Montrer que

$$p = \sum_{\ell=1}^L \mathbb{P}[\langle \mathbf{X}, u \rangle \in D_\ell] \mathbb{E} \left[\max \left\{ \frac{1}{d} \sum_{i=1}^d \exp[\sigma(Y_\ell u_i + Z_i - \langle \mathbf{Z}, u \rangle u_i)] - K, 0 \right\} \right]. \tag{1}$$

- 10. (a) Proposer une partition (D_1, \dots, D_ℓ) telle que pour $\ell \in \llbracket 1, L \rrbracket$, $\mathbb{P}[\langle \mathbf{X}, u \rangle \in D_\ell] = 1/L$.
- (b) (♠) On prend $L = 40$. Partant de (1), donner une estimation de p , noté $\hat{p}_n(u)$ avec allocation proportionnelle pour chacune de ces strates D_ℓ , $\ell \in \llbracket 1, L \rrbracket$. Donner une estimation de la variance de la méthode.
Indication. On pourra considérer comme exacte l'approximation polynomiale de l'inverse généralisé de la loi $\mathcal{N}(0, 1)$.
- 11. (♠) Déterminer l'efficacité relative de ces deux méthodes d'estimation de p . Commenter le résultat.

Auto-évaluation. Renseigner la grille suivante avec **T** le temps de calcul médian pour 100 répétitions du code et le nombre de fois où les différents éléments cités sont utilisés dans votre code.

Exercice 1.

Référence	T	c()	for	while	if-else	Vectorize	apply
$\hat{\delta}_n$	627 ms	0	0	1	0	0	0
$\hat{\Delta}_n(\theta)$	857 ms	0	0	1	0	0	0
Code	T	c()	for	while	if-else	Vectorize	apply
$\hat{\delta}_n$							
$\hat{\Delta}_n(\theta)$							

Exercice 2.

Référence	T	c()	for	while	if-else	Vectorize	apply
$\hat{\delta}_n$	7.76 ms	0	0	0	0	0	0
$\hat{\delta}_n(A)$	10.5 ms	0	0	0	0	0	0
Code	T	c()	for	while	if-else	Vectorize	apply
$\hat{\delta}_n$							
$\hat{\delta}_n(A)$							

Référence	T	c()	for	while	if-else	Vectorize	apply
$\hat{\mathcal{I}}_n$	1.95 ms	0	0	0	0	0	0
$\hat{\mathcal{I}}_n(b)$ (★)	2.23 ms	0	0	0	0	0	0
$\hat{\mathcal{I}}_n(b)$ (★★)	12.7 ms	0	1	0	0	0	0
$\hat{\mathcal{I}}_n(b)$ (★★★)	2.51 ms	0	1	0	0	0	0

Code	T	c()	for	while	if-else	Vectorize	apply
$\hat{\mathcal{I}}_n$							
$\hat{\mathcal{I}}_n(b)$							

Référence	T	c()	for	while	if-else	Vectorize	apply
\hat{p}_n	78.1 ms	0	0	0	0	0	0
$\hat{p}_n(u)$	115.1 ms	0	1	0	0	0	0

Code	T	c()	for	while	if-else	Vectorize	apply
\hat{p}_n							
$\hat{p}_n(u)$							