

Cours et travail dirigé autour de la méthode de Monte-Carlo

La première partie de ce document doit être considérée comme du cours.

Principe de la méthode de Monte-Carlo. On considère une variable aléatoire $X : \Omega \rightarrow \mathbb{R}^d$ et une fonction mesurable $f : \mathbb{R}^d \rightarrow \mathbb{R}$ telle que

$$\mathbb{E}[|f(X)|] < \infty.$$

Le but de la méthode de Monte-Carlo est de calculer de façon approchée $\mathbb{E}[f(X)]$. Pour cela, on suppose que l'on sait simuler la variable aléatoire X , c'est à dire générer une suite i.i.d. $(X_i)_{i \geq 1}$ telle que $X_i \stackrel{\mathcal{L}}{=} X$. L'idée est d'approcher $\mathbb{E}[f(X)]$ par

$$M_n = \frac{1}{n} \sum_{i=1}^n f(X_i), \quad (1)$$

avec n assez grand. En effet, la loi forte des grands nombres assure que $M_n \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[f(X)]$, p.s. Cela nous dit que pour n assez grand, M_n sera proche de $\mathbb{E}[f(X)]$. En revanche, nous ne savons pas à ce stade quel n choisir pour obtenir une précision ε donnée. Pour cela, nous faisons l'hypothèse supplémentaire que $\mathbb{E}[f(X)^2] < \infty$, et on pose

$$\sigma^2 = \mathbb{E}[f(X)^2] - \mathbb{E}[f(X)]^2.$$

Alors, le théorème de la limite centrale assure que

$$\frac{\sqrt{n}}{\sigma} (M_n - \mathbb{E}[f(X)]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (2)$$

On fait alors l'approximation que pour n suffisamment grand,

$$\frac{\sqrt{n}}{\sigma} (M_n - \mathbb{E}[f(X)]) \stackrel{\mathcal{L}}{\approx} \mathcal{N}(0, 1).$$

Comme pour $G \sim \mathcal{N}(0, 1)$, $\mathbb{P}(-1,96 \leq G \leq 1,96) \approx 0,95$, on a qu'avec probabilité 95%,

$$\mathbb{E}[f(X)] \in \left[M_n - \frac{1,96\sigma}{\sqrt{n}}, M_n + \frac{1,96\sigma}{\sqrt{n}} \right].$$

L'intervalle $\left[M_n - \frac{1,96\sigma}{\sqrt{n}}, M_n + \frac{1,96\sigma}{\sqrt{n}} \right]$ est appelé intervalle de confiance à 95% pour $\mathbb{E}[f(X)]$. De même, comme $\mathbb{P}(-2,58 \leq G \leq 2,58) \approx 0,99$, l'intervalle $\left[M_n - \frac{2,58\sigma}{\sqrt{n}}, M_n + \frac{2,58\sigma}{\sqrt{n}} \right]$ est un intervalle de confiance à 99% pour $\mathbb{E}[f(X)]$.

En pratique, lorsque l'on cherche à calculer $\mathbb{E}[f(X)]$ de façon approchée, il est rare de connaître la valeur exacte de σ . Parfois, on connaît a priori un majorant de σ - noté ici $\bar{\sigma}$ - et dans ce cas, on sait qu'avec une probabilité au moins égale à 95%, $\mathbb{E}[f(X)]$ appartient à $\left[M_n - \frac{1,96\bar{\sigma}}{\sqrt{n}}, M_n + \frac{1,96\bar{\sigma}}{\sqrt{n}} \right]$,

et cet intervalle est donc à nouveau un intervalle de confiance à 95%. Néanmoins, dans la grande majorité des cas, on ne sait rien sur σ et on cherche à l'estimer également par Monte-Carlo. Ainsi,

$$V_n = \frac{1}{n} \sum_{i=1}^n f(X_i)^2 - M_n^2$$

converge presque sûrement vers σ^2 par la loi forte des grands nombres. Grâce à (2) et au Théorème de Slutsky, on en déduit que

$$\frac{\sqrt{n}}{\sqrt{V_n}}(M_n - \mathbb{E}[f(X)]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

En répétant le même raisonnement que précédemment, on en déduit que

$$\left[M_n - \frac{1,96\sqrt{V_n}}{\sqrt{n}}, M_n + \frac{1,96\sqrt{V_n}}{\sqrt{n}} \right]$$

est un intervalle de confiance à 95% pour $\mathbb{E}[f(X)]$.

1 Calcul d'intégrale : sommes de Riemann ou méthode de Monte-Carlo ?

Dans cette première partie, nous étudions l'approximation numérique de l'intégrale $I = \int_{[0,1]^d} f(x)dx$, où $f : [0, 1]^d \rightarrow \mathbb{R}$ est une fonction mesurable intégrable. Le but est de comparer la méthode usuelle des sommes de Riemann avec la méthode de Monte-Carlo. Nous commençons par le cas de la dimension $d = 1$.

1.1 En dimension $d = 1$

En dimension 1, les sommes de Riemann s'écrivent pour $n \in \mathbb{N}^*$,

$$S_n = \frac{1}{n} \sum_{i=0}^{n-1} f\left(\frac{i}{n}\right).$$

1. On suppose $f : [0, 1] \rightarrow \mathbb{R}$ continue. On rappelle le théorème de Heine : toute application continue sur un compact est uniformément continue. Ainsi, f satisfait

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x, y \in [0, 1], |x - y| \leq \eta \implies |f(x) - f(y)| \leq \varepsilon.$$

En déduire que $S_n \xrightarrow{n \rightarrow +\infty} I$.

En fait, on peut montrer que $S_n \xrightarrow{n \rightarrow +\infty} I$ pour une classe un peu plus grande de fonctions, les fonctions réglées qui incluent notamment les fonctions f continues par morceaux. En revanche, on ne dispose pas de vitesse de convergence à ce stade. Pour cela, on a besoin de faire des hypothèses de régularité sur f . La plus usuelle est de supposer f lipschitzienne, i.e.

$$\exists K > 0, \forall x, y \in [0, 1], |f(x) - f(y)| \leq K|x - y|.$$

2. On suppose f Lipschitzienne. Montrer que

$$|S_n - I| \leq \frac{K}{2n}.$$

En déduire que pour calculer I avec une précision de $\varepsilon > 0$, il faut un temps de calcul proportionnel à $\frac{K}{\varepsilon}$.

3. On suppose désormais f telle que $\int_0^1 f(u)^2 du < \infty$, et on note $\sigma^2 = \int_0^1 f(u)^2 du - \left(\int_0^1 f(u) du\right)^2$. Ecrire une méthode de Monte-Carlo qui permet d'approcher I . Donner l'intervalle de confiance à 95% en supposant σ connu. En déduire que pour obtenir une précision de $\varepsilon > 0$, il faut un temps de calcul proportionnel à $\frac{\sigma^2}{\varepsilon^2}$.

▷ **A retenir** : En dimension 1, sauf éventuellement dans le cas où f n'est pas régulière, il faut préférer les sommes de Riemann (temps de calcul en $O(\varepsilon^{-1})$) à la méthode de Monte-Carlo (temps de calcul en $O(\varepsilon^{-2})$). Si f est plus régulière, on préférera utiliser la méthode des trapèzes ($f \in \mathcal{C}^2$) ou de Simpson ($f \in \mathcal{C}^4$) pour obtenir des temps de calcul respectivement en $O(\varepsilon^{-1/2})$ et $O(\varepsilon^{-1/4})$.

1.2 En dimension d quelconque

Pour $d \geq 1$, les sommes de Riemann s'écrivent

$$S_n = \frac{1}{n^d} \sum_{i_1=0}^{n-1} \cdots \sum_{i_d=0}^{n-1} f\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right).$$

4. Montrer que

$$S_n - I = \sum_{i_1=0}^{n-1} \cdots \sum_{i_d=0}^{n-1} \int_{\frac{i_1}{n}}^{\frac{i_1+1}{n}} \cdots \int_{\frac{i_d}{n}}^{\frac{i_d+1}{n}} f(x_1, \dots, x_d) - f\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right) dx_1 \cdots dx_d.$$

5. On suppose $f : [0, 1]^d \rightarrow \mathbb{R}$ lipschitzienne, i.e.

$$\forall x, y \in [0, 1]^d, |f(x) - f(y)| \leq K \|x - y\|_1.$$

Montrer que $|S_n - I| \leq \frac{Kd}{2n}$. En déduire que pour obtenir une précision de $\varepsilon > 0$, il faut un temps de calcul proportionnel à $\left(\frac{Kd}{2}\right)^d \frac{1}{\varepsilon^d}$.

6. On suppose désormais f telle que $\int_{[0,1]^d} f(x)^2 dx < \infty$, et on pose $\sigma^2 = \int_{[0,1]^d} f(x)^2 dx - \left(\int_{[0,1]^d} f(x) dx\right)^2$. Ecrire une méthode de Monte-Carlo qui permet d'approcher I et donner l'intervalle de confiance à 95% avec σ connu. En déduire qu'il faut un temps de calcul proportionnel à $\frac{d\sigma^2}{\varepsilon^2}$ pour une précision de ε .

▷ **A retenir** : En pratique, la dimension d et la fonction f sont imposées par le problème que l'on considère. Pour une précision de ε , le temps demandé par le calcul des sommes de Riemann est en $O(\varepsilon^{-d})$ lorsque $\varepsilon \rightarrow 0$. On parle de "malédiction de la dimension" : plus la dimension est grande, plus le calcul est gourmand en temps. En revanche, pour la méthode de Monte-Carlo, le temps de calcul est en $O(\varepsilon^{-2})$, quelle que soit la dimension. Bien sûr, lorsque f est régulière, on peut améliorer les sommes de Riemann avec des méthodes de type trapèzes mais qui sont frappées aussi par la malédiction de la dimension. Pour fixer les idées (car il faudrait nuancer selon les cas), la méthode de Monte-Carlo s'avère être la plus efficace à partir de $d = 4$ ou 5.

2 Quelques exemples d'applications de la méthode de Monte-Carlo

L'idée de la méthode de Monte-Carlo est quelque part aussi ancienne que les probabilités et leur interprétation fréquentiste. Par exemple, dans un jeu de pile ou face, si on veut s'assurer que le jeu est équitable, on va lancer un grand nombre de fois n la pièce, et on va vérifier que le ratio

n_F/n (où n_F est le nombre de fois où on a obtenu “Face”) est proche de $1/2$. Plus précisément si $1/2$ appartient à l’intervalle de confiance à 95%

$$\left[\frac{n_F}{n} - 1,96 \frac{\sqrt{\frac{n_F}{n} \left(1 - \frac{n_F}{n}\right)}}{\sqrt{n}}, \frac{n_F}{n} + 1,96 \frac{\sqrt{\frac{n_F}{n} \left(1 - \frac{n_F}{n}\right)}}{\sqrt{n}} \right],$$

on aura tendance à accepter l’hypothèse que le jeu est équitable. En revanche, si $1/2$ n’est pas dans cet intervalle, on privilégiera l’hypothèse que la pièce est biaisée¹. La première utilisation de la méthode de Monte-Carlo pour faire un calcul numérique est souvent attribuée à Buffon. Il a proposé en 1733 une expérience qui consiste à lancer des aiguilles sur un parquet dont la longueur est inférieure à la largeur des lattes du parquet. Sous l’hypothèse d’un lancer uniforme des aiguilles, on peut montrer la probabilité qu’une aiguille croise une rainure du parquet est inversement proportionnelle à π . On peut donc estimer π en lançant un grand nombre de fois n les aiguilles et en calculant le ratio n_c/n , où n_c est le nombre de fois où l’aiguille était à cheval sur deux lattes.

Exercice 1. (*L’aiguille de Buffon*) On note ℓ la longueur de l’aiguille et L la largeur des planches du parquet, et on suppose $\ell \leq L$. On définit l’axe des abscisses par une droite perpendiculaire aux planches, l’origine des abscisses étant une intersection avec une rainure. On note X l’abscisse minimale de l’aiguille et Θ l’angle non orienté entre l’aiguille et l’axe des abscisses. On suppose que, modulo L , $X \sim \mathcal{U}([0, L])$ et est indépendante de l’angle $\Theta \sim \mathcal{U}([0, \pi/2])$. Montrer que la probabilité que l’aiguille croise une rainure du parquet est égale à $\frac{2\ell}{\pi L}$. Construire un intervalle de confiance à 95% pour π , exprimé à l’aide de n_c et n .

Aujourd’hui, la méthode de Monte-Carlo est utilisée dans de très nombreux domaines, et il serait impossible d’être exhaustif et nous donnons seulement quelques exemples ici. Tout d’abord, elle est utilisée pour remplacer des simulations réelles qui sont souvent plus coûteuses. Par exemple, en biologie, il peut être coûteux en temps et en argent de tester l’effet d’un principe actif sur des cellules. Si on est capable de bien modéliser les réactions chimiques à l’intérieur des cellules, on peut remplacer des tests par des simulations et la méthode de Monte-Carlo pour calculer par exemple la probabilité qu’un principe actif guérisse d’une maladie ou bien la quantité minimum de médicament à prendre pour qu’il fasse effet. Plusieurs projets développent aujourd’hui des “simulateurs de cellules”, voir par exemple <http://www.e-cell.org/> ou <http://www.nrcam.uchc.edu/>. De façon générale, les méthodes de simulations et de calcul par Monte-Carlo peuvent être utiles à chaque fois que la simulation réelle est coûteuse. Cela est le cas pour simuler des réactions nucléaires, pour tester la résistance de certains matériaux à des stress importants, pour dimensionner un ouvrage d’art etc. En météorologie, la méthode de Monte-Carlo permet de donner des prévisions avec un indice de confiance <http://www.meteofrance.com/prevision-probabiliste>. Les banques et les sociétés d’assurances doivent régulièrement montrer aux autorités de contrôle qu’elles ont suffisamment de fonds propres pour faire face à un événement extrême (krach boursier, faillite d’une entreprise, catastrophe naturelle, etc.) Pour ce faire, elles doivent avoir des modèles d’évolutions de leurs portefeuilles, faire des simulations et prouver que par exemple dans 99% des cas, elles sont capables de résister. En finance toujours, la théorie développée par Black, Merton et Scholes au début des années 1970 fait apparaître le prix de produits dérivés comme une espérance, qui peut donc être évalué à l’aide de la méthode de Monte-Carlo. On présente ci-dessous un exemple concret de calcul de prix.

1. L’utilisation de tests statistiques permet de donner un cadre général plus formel pour tester si une hypothèse est vraie ou non.

Exemple de calcul par Monte-Carlo. On suppose qu'une banque souhaite proposer à ses clients un produit qui rémunère la meilleure performance des actions du CAC 40 sur une période de un an, sauf si celle-ci est supérieure à 10% en quel cas elle rémunère à 10%. La question est de savoir à quel prix les clients doivent acheter ce produit. On modélise le ratio entre le prix des actifs au bout d'un an et celui d'aujourd'hui par $S_i = \exp\left(-\frac{\sigma^2}{2} + \sigma W_i\right)$, $i = 1, \dots, 40$, avec

$$W \sim \mathcal{N}_{40}(0, (1 - \rho)I + \rho J),$$

I la matrice identité, et J la matrice dont tous les éléments sont égaux à 1. Ces deux matrices sont carrées de dimension 40. Le coefficient $\rho \in [0, 1]$ paramètre la dépendance entre les actifs, et on a fait ici pour simplifier l'hypothèse que tous les actifs ont la même volatilité $\sigma > 0$. Le produit que souhaite proposer la banque paye $\min(\max_{i=1, \dots, 40} S_i, \frac{11}{10})$, et on admettra que son prix théorique est donné par

$$P = \mathbb{E} \left[\min \left(\max_{i=1, \dots, 40} S_i, \frac{11}{10} \right) \right].$$

Exercice 2. Soit $G \sim \mathcal{N}_{40}(0, I)$ et $N \sim \mathcal{N}(0, 1)$ indépendante de G . On définit $G'_i = \sqrt{\rho}N + \sqrt{1 - \rho}G_i$. Montrer que G' et W ont même loi.

Le calcul de P revient à faire une intégration en dimension 40 (même 41 si on utilise le résultat de l'exercice). Il est donc souhaitable d'utiliser la méthode de Monte-Carlo, et nous donnons ci-dessous le code Python permettant de calculer P . Une exécution de ce code nous indique que $[1.06833159304, 1.07134023987]$ est un intervalle de confiance à 95% pour P .

```
import random
import math

K=40; #Nombre d'actifs
N=10000; #Nombre de simulations

sig=0.3; #Volatilité
rho=0.7; #Corrélation/Dépendance entre actifs

M=0;
V=0;
srho=math.sqrt(rho);
srho2=math.sqrt(1-rho);

for i in range(N):
    z=random.gauss(0.0,1.0)
    x=0
    for j in range(K):
        x=max(x,math.exp(-0.5*sig**2+sig*(srho*z+srho2*random.gauss(0.0,1.0))))
        x=min(x,1.1)
    M=M+x/N
    V=V+x**2/N
print "Intervalle de confiance à 95% : [", M-1.96*math.sqrt((V-M**2)/N), ", ", M+1.96*math.sqrt((V-M**2)/N), "]"
```

3 Améliorations de la méthode de Monte-Carlo : méthodes de réduction de variance

La méthode de Monte-Carlo permet de calculer des intégrales avec une vitesse de convergence qui ne dépend pas de la dimension. Cette vitesse de convergence est en σ/\sqrt{n} . La décroissance en $1/\sqrt{n}$ est celle du Théorème de la limite centrale, et on ne peut pas la changer sauf si on utilise une autre méthode que la méthode de Monte-Carlo. Par conséquent, la seule chose que l'on peut espérer améliorer est la variance σ^2 : par exemple une diminution d'un facteur 10 de σ^2 permet d'obtenir la même précision avec 10 fois moins de simulations. Nous allons donner ici un aperçu des méthodes génériques pour réduire la variance de l'estimateur de Monte-Carlo. Par souci de simplicité, nous présentons ces méthodes en dimension $d = 1$. Elles se généralisent sans encombre à une dimension quelconque, et c'est là qu'elles sont pertinentes pour des applications concrètes.

3.1 Méthode de fonction d'importance

Pour fixer les idées, on considère que X est une variable aléatoire réelle de densité p . On souhaite calculer $\mathbb{E}[f(X)]$ par la méthode de Monte-Carlo. Soit Y une variable aléatoire réelle de densité q strictement positive. On a alors,

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)p(x)dx = \int_{\mathbb{R}} f(x)\frac{p(x)}{q(x)}q(x)dx = \mathbb{E}\left[f(Y)\frac{p(Y)}{q(Y)}\right].$$

On suppose que :

- on sait simuler les variables aléatoires X et Y ,
- les densités p et q sont connues et calculables.

Soient $(Y_i)_{i \geq 1}$ une suite i.i.d telle que $Y_i \stackrel{\mathcal{L}}{=} Y$. Alors, on peut approcher $\mathbb{E}[f(X)]$ par M_n défini en (1) ou bien par

$$\tilde{M}_n = \frac{1}{n} \sum_{i=1}^n f(Y_i) \frac{p(Y_i)}{q(Y_i)}.$$

On préférera \tilde{M}_n à M_n si $\tilde{\sigma}^2 = \mathbb{E}\left[\left(f(Y)\frac{p(Y)}{q(Y)}\right)^2\right] - \mathbb{E}[f(X)]^2 < \sigma^2$. Le choix de la loi de Y est donc crucial pour mettre en oeuvre cette méthode.

Exercice 3. Soit $g : [0, 1] \rightarrow \mathbb{R}$ mesurable de carré intégrable. On suppose que l'on cherche à calculer $I = \int_0^1 g(x)e^{-x}dx$ par la méthode de Monte-Carlo. On se donne une suite i.i.d. $(U_i)_{i \geq 1}$ de variables aléatoires uniformes sur $[0, 1]$. Rappeler la loi de $-\log(U_1)$. Montrer que $M_n = \frac{1}{n} \sum_{i=1}^n g(U_i)e^{-U_i}$ et $\tilde{M}_n = \frac{1}{n} \sum_{i=1}^n g(-\log(U_i))\mathbf{1}_{\{-\log(U_i) \leq 1\}}$ convergent presque sûrement vers I . Lequel de ces deux estimateurs préférer ?

3.2 Méthode des variables antithétiques

Cette méthode utilise des symétries de la loi que l'on simule pour avoir “gratuitement” de nouveaux tirages. Par exemple, on sait que pour $U \sim \mathcal{U}([0, 1])$, $1 - U$ suit également la loi $\mathcal{U}([0, 1])$. Prenons l'exemple où on souhaite calculer $I = \int_0^1 f(x)dx$, avec f mesurable de carré intégrable. La méthode de Monte-Carlo usuelle consiste à utiliser

$$M_n = \frac{1}{n} \sum_{i=1}^n f(U_i),$$

où $(U_i)_{i \geq 1}$ est une suite i.i.d. de variables aléatoires uniformes sur $[0, 1]$. La méthode des variables antithétiques consiste à utiliser :

$$\tilde{M}_n = \frac{1}{n} \sum_{i=1}^n \frac{f(U_i) + f(1 - U_i)}{2}$$

On a $\mathbf{Var}\left(\frac{f(U)+f(1-U)}{2}\right) = \frac{1}{2}[\mathbf{Var}(f(U)) + \mathbf{Cov}(f(U), f(1-U))] \leq \mathbf{Var}(f(U))$, et donc \tilde{M}_n converge toujours mieux que M_n . Cependant, si dans le calcul numérique de M_n et \tilde{M}_n , le temps de calcul pour évaluer f est bien plus grand que celui nécessaire à générer un nombre aléatoire, il est plus judicieux de comparer \tilde{M}_n avec M_{2n} . En effet ces deux estimateurs demandent tous les deux $2n$ calculs de la fonction f . On a

$$\begin{aligned}\mathbf{Var}(M_{2n}) &= \frac{1}{2n} \mathbf{Var}(f(U)), \\ \mathbf{Var}(\tilde{M}_n) &= \frac{1}{2n} [\mathbf{Var}(f(U)) + \mathbf{Cov}(f(U), f(1-U))].\end{aligned}$$

Par conséquent, on préférera \tilde{M}_n à M_{2n} si

$$\mathbf{Cov}(f(U), f(1-U)) \leq 0.$$

C'est généralement ce critère qui est utilisé pour décider si on utilise ou pas la méthode des variables antithétiques.

Exercice 4. On suppose $g, h : [0, 1] \rightarrow \mathbb{R}$ mesurables et de carré intégrable. On suppose g croissante et h décroissante. Montrer que $\mathbf{Cov}(g(U), h(U)) \leq 0$ (On pourra calculer $\mathbb{E}[(g(U_1) - g(U_2))(h(U_1) - h(U_2))]$). En déduire que si f est monotone,

$$\mathbf{Cov}(f(U), f(1-U)) \leq 0.$$

3.3 Méthode des variables de contrôle

On considère le cadre donné en introduction avec une variable aléatoire réelle X et une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\mathbb{E}[f(X)^2] < \infty$. On cherche à calculer $\mathbb{E}[f(X)]$ par méthode de Monte-Carlo. On suppose que l'on connaît une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ de carré intégrable telle que $\mathbb{E}[g(X)]$ est connu explicitement et telle que $\mathbf{Var}(g(X)) > 0$. On considère alors, pour $\lambda \in \mathbb{R}$,

$$M_n^\lambda = \frac{1}{n} \sum_{i=1}^n (f(X_i) + \lambda g(X_i)) - \lambda \mathbb{E}[g(X)].$$

Par la loi forte des grands nombres, M_n^λ converge presque sûrement vers $\mathbb{E}[f(X)]$. Le théorème de la limite centrale donne la convergence en loi de $\frac{\sqrt{n}}{\sigma_\lambda} (M_n^\lambda - \mathbb{E}[f(X)])$ vers $\mathcal{N}(0, 1)$, où

$$\begin{aligned}\sigma_\lambda^2 &= \mathbb{E}[(f(X) + \lambda g(X))^2] - \mathbb{E}[f(X) + \lambda g(X)]^2 \\ &= \mathbf{Var}(f(X)) + 2\lambda \mathbf{Cov}(f(X), g(X)) + \lambda^2 \mathbf{Var}(g(X)).\end{aligned}$$

Le meilleur choix de λ est $\lambda^* = -\frac{\mathbf{Cov}(f(X), g(X))}{\mathbf{Var}(g(X))}$, pour lequel σ_λ^2 atteint son minimum

$$\underline{\sigma}^2 = \frac{\mathbf{Var}(f(X))\mathbf{Var}(g(X)) - \mathbf{Cov}(f(X), g(X))^2}{\mathbf{Var}(g(X))}.$$

On a toujours $\underline{\sigma}^2 \leq \sigma^2$ par construction. Le problème pour appliquer cette méthode est qu'il faut connaître λ^* ce qui n'est a priori pas le cas. La solution est de l'estimer aussi par Monte-Carlo. Pour $h : \mathbb{R} \rightarrow \mathbb{R}$, on introduit la notation

$$M_n(h) := \frac{1}{n} \sum_{i=1}^n h(X_i).$$

On a alors par la loi forte des grands nombres :

$$\frac{M_n(fg) - M_n(f)M_n(g)}{M_n(g^2) - M_n(g)^2} \xrightarrow{n \rightarrow +\infty} \frac{\mathbf{Cov}(f(X), g(X))}{\mathbf{Var}(g(X))}, \text{ p.s.}$$

On définit alors

$$\tilde{M}_n = M_n(f) - \frac{M_n(fg) - M_n(f)M_n(g)}{M_n(g^2) - M_n(g)^2} (M_n(g) - \mathbb{E}[g(X)]).$$

En utilisant le théorème de Slutsky, on peut montrer (exercice !) que $\frac{\sqrt{n}}{\sigma}(\tilde{M}_n - \mathbb{E}[f(X)]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$, puis que

$$\frac{\sqrt{n}}{\sqrt{\tilde{V}_n}}(\tilde{M}_n - \mathbb{E}[f(X)]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

avec $\tilde{V}_n = M_n(f^2) - M_n(f)^2 - \frac{(M_n(fg) - M_n(f)M_n(g))^2}{M_n(g^2) - M_n(g)^2}$. En particulier,

$$\left[\tilde{M}_n - \frac{1,96\sqrt{\tilde{V}_n}}{\sqrt{n}}, \tilde{M}_n + \frac{1,96\sqrt{\tilde{V}_n}}{\sqrt{n}} \right]$$

est un intervalle de confiance à 95% pour $\mathbb{E}[f(X)]$.

La variable aléatoire $g(X)$ est appelée variable de contrôle. On peut généraliser cette méthode en considérant plusieurs variables de contrôle $g_1(X), \dots, g_n(X)$ dont on connaît explicitement l'espérance.

Exercice 5. On considère un ensemble $A \subset [0, 1]^2$ tel que $A' = \{(u, v) \in [0, 1]^2, u + v \leq 1\} \subset A$. Par exemple, A peut être le quart de disque $\{(u, v) \in [0, 1]^2, u^2 + v^2 \leq 1\}$ ou plus généralement l'ensemble $\{(u, v) \in [0, 1]^2, u^\alpha + v^\beta \leq 1\}$ avec $\alpha, \beta \geq 1$. On cherche à calculer l'aire de A par méthode de Monte-Carlo, que l'on note $|A|$.

1. On considère $(U_i, V_i)_{i \geq 1}$ une suite i.i.d., avec $U_1, V_1 \sim \mathcal{U}([0, 1])$ et indépendantes. Montrer que $M_n = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_A(U_i, V_i)$ converge p.s. vers $|A|$, puis que $\frac{\sqrt{n}}{\sqrt{M_n(1-M_n)}}(M_n - |A|) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$. Construire un intervalle de confiance à 95% pour $|A|$.
2. Calculer $\mathbb{E}[\mathbf{1}_{A'}(U_1, V_1)]$ et $\mathbf{Var}(\mathbf{1}_{A'}(U_1, V_1))$. Donner en fonction de $|A|$ la valeur de $\lambda \in \mathbb{R}$ qui minimise $\sigma_\lambda^2 = \mathbb{E}[(\mathbf{1}_A(U_1, V_1) + \lambda \mathbf{1}_{A'}(U_1, V_1))^2] - \mathbb{E}[\mathbf{1}_A(U_1, V_1) + \lambda \mathbf{1}_{A'}(U_1, V_1)]^2$ ainsi que la valeur minimale $\underline{\sigma}^2 = \min_{\lambda \in \mathbb{R}} \sigma_\lambda^2$.
Vérifier que $\underline{\sigma}^2 = 0$ lorsque $|A| = 1/2$. Expliquer pourquoi ce résultat est attendu.
3. On note $\tilde{M}_n = M_n + 2(M_n - 1) \left[\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{A'}(U_i, V_i) - 1/2 \right] = 1 + 2(M_n - 1) \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{A'}(U_i, V_i)$. Construire un intervalle de confiance à 95% à partir de \tilde{M}_n .

3.4 Méthode de stratification

Le principe de cette méthode est de chercher à effectuer plus de simulations “là où c'est le plus utile”. Pour l'illustrer nous considérons le cas du calcul de $I = \int_0^1 f(u)du$ avec f mesurable, de carré intégrable. On découpe l'intervalle $[0, 1]$ en deux sous-intervalles (strates) $[0, a]$ et $[a, 1]$ avec $0 < a < 1$. On se donne également un nombre $\alpha \in]0, 1[$, et on va effectuer une proportion α des simulations dans $[0, a]$ et une proportion $1 - \alpha$ des simulations dans $[a, 1]$. Concrètement, cela signifie que l'on va utiliser

$$\tilde{M}_n = \frac{a}{\alpha n} \sum_{i=1}^{\lfloor \alpha n \rfloor} f(aU_i) + \frac{1-a}{(1-\alpha)n} \sum_{i=\lfloor \alpha n \rfloor + 1}^n f(a + (1-a)U_i),$$

où $\lfloor x \rfloor$ désigne la partie entière de x . Par commodité pour l'étude mathématique, on préfère travailler avec deux suites indépendantes et i.i.d. $(U_i)_{i \geq 1}$ et $(V_i)_{i \geq 1}$ de variables aléatoires uniformes sur $[0, 1]$, et on continue de noter

$$\tilde{M}_n = \frac{a}{\alpha n} \sum_{i=1}^{\lfloor \alpha n \rfloor} f(aU_i) + \frac{1-a}{(1-\alpha)n} \sum_{i=1}^{n-\lfloor \alpha n \rfloor} f(a + (1-a)V_i).$$

La loi forte des grands nombre donne la convergence presque sûre de \tilde{M}_n vers

$$a \int_0^1 f(au)du + (1-a) \int_0^1 f(a + (1-a)u)du = I.$$

On s'intéresse à la variance de l'estimateur \tilde{M}_n :

$$\begin{aligned} \mathbf{Var}(\tilde{M}_n) &= \frac{a^2}{(\alpha n)^2} \lfloor \alpha n \rfloor \mathbf{Var}(f(aU)) + \frac{(1-a)^2}{((1-\alpha)n)^2} (n - \lfloor \alpha n \rfloor) \mathbf{Var}(f(a + (1-a)U)) \\ &\underset{n \rightarrow +\infty}{\sim} \frac{1}{n} \sigma_{\alpha,a}^2, \end{aligned}$$

avec $\sigma_{\alpha,a}^2 = \frac{1}{\alpha} [a \int_0^a f(u)^2 du - (\int_0^a f(u) du)^2] + \frac{1}{1-\alpha} [(1-a) \int_a^1 f(u)^2 du - (\int_a^1 f(u) du)^2]$. Si $\sigma_{\alpha,a}^2 < \sigma^2$, on aura effectivement réduit la variance.

Exercice 6. Montrer que $\frac{\sqrt{n}}{\sigma_{\alpha,a}}(\tilde{M}_n - I) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$.

Bien évidemment, on souhaite si possible connaître une valeur de (α, a) qui minimise $\sigma_{\alpha,a}^2$. La minimisation en α est simple, elle est obtenue pour

$$\alpha^* = \frac{\sqrt{\frac{a \int_0^a f(u)^2 du - (\int_0^a f(u) du)^2}{(1-a) \int_a^1 f(u)^2 du - (\int_a^1 f(u) du)^2}}}{1 + \sqrt{\frac{a \int_0^a f(u)^2 du - (\int_0^a f(u) du)^2}{(1-a) \int_a^1 f(u)^2 du - (\int_a^1 f(u) du)^2}}},$$

en on a

$$\sigma_{\alpha^*,a}^2 = \left(\sqrt{a \int_0^a f(u)^2 du - \left(\int_0^a f(u) du \right)^2} + \sqrt{(1-a) \int_a^1 f(u)^2 du - \left(\int_a^1 f(u) du \right)^2} \right)^2.$$

En revanche, on ne connaît généralement pas a priori la valeur optimale de a . Pour cela, il faut mettre en oeuvre des méthodes de Monte-Carlo dites adaptatives qui effectuent la minimisation au fur et à mesure des simulations. Ces méthodes dépassent largement le cadre de ce cours. Enfin, mentionnons ici que cette méthode se généralise facilement à n strates (sur $[0, 1]$, en considérant les intervalles $[0, a_1]$, $[a_1, a_2]$, ..., $[a_n, 1]$ avec $0 < a_1 < \dots < a_n < 1$).

Exercice 7. Dans cette exercice, on ne cherche pas à optimiser la strate a et on fixe $a = 1/2$.

1. Montrer que $\mathbf{Var}(f(U/2)) \leq \mathbf{Var}(f(1/2 + U/2)) \iff \alpha^* \leq 1/2$. Interpréter ce résultat.
2. On s'intéresse au cas $f(u) = |1-2u|$. Montrer que $\alpha^* = 1/2$, puis calculer $\sigma_{1/2,1/2}^2$. Comparer à σ^2 et commenter.
3. On considère maintenant $f(u) = 1 - 2u$ si $u \in [0, 1/2]$ et $f(u) = 2(2u - 1)$ si $u \in [1/2, 1]$. Donner la valeur de α^* et de $\frac{\sigma_{1/2,1/2}^2}{\sigma^2}$ dans ce cas.