

UNIVERSITÉ PARIS - DAUPHINE



---

## Projet Monte - Carlo

---

Shana ZIRAH  
Eva ZAGURY

*Tuteur* : Julien STOEHR

27 décembre 2020

# Table des matières

<b>1</b>	<b>Exercice 1</b>	<b>2</b>
1.1	Méthode n°1 - Estimation de $a$ . . . . .	4
1.2	Méthode n°2 - Estimateur ponctuel . . . . .	9
<b>2</b>	<b>Exercice 2</b>	<b>11</b>
<b>3</b>	<b>Exercice 3</b>	<b>14</b>

Note : Toutes les valeurs affichées dans ce projet sont implémentées avec une graine de 111.

## 1 Exercice 1

### Simulation suivant la densité f

1. Pour simuler suivant la densité f, on utilise l'algorithme du rejet. L'algorithme du rejet nous indique qu'à partir d'une densité g de  $\mathbb{R}^d$  pour laquelle on sait simuler des variables aléatoires et d'une constante  $M \geq 1$  satisfaisant :

$$f(x) \leq Mg(x), \text{ pour tout } x \text{ dans } \mathbb{R}^d$$

il est possible de simuler une variable aléatoire  $X = (X_1, \dots, X_d)$  de densité f de  $\mathbb{R}^d$ .

Ici, f est définie pour  $(x, y) \in \mathbb{R}^d$  par  $f(x, y) = a\psi(x, y)$  avec  $a > 0$ . D'où,

$$f(x, y) \leq Mg(x, y), \forall (x, y) \in \mathbb{R}^2 \Leftrightarrow a\psi(x, y) \leq Mg(x, y) \forall x \in \mathbb{R}^2$$

Or,  $a > 0$ , donc on peut poser  $m := \frac{M}{a}$  afin d'obtenir

$$\forall (x, y) \in \mathbb{R}^2, \psi(x, y) \leq mg(x, y)$$

Le temps d'attente T défini tel que  $Y_T$  suit la loi de densité f est

$$T := \inf\{n \geq 1 : U_n \leq \rho(Y_{1,n}, Y_{2,n})\}$$

où  $(Y_{1,n}, Y_{2,n})$  sont des réalisations suivant la densité g

Par conséquent, pour obtenir une réalisation suivant f, il suffit de simuler  $U \sim \mathcal{U}(0, 1)$  et  $(Y_1, Y_2)$  un vecteur aléatoire de densité g jusqu'à ce que  $Umg(Y_1, Y_2) \leq \psi(Y_1, Y_2)$  ou encore  $U \leq \rho(Y_1, Y_2)$ .

2. Le choix de m et g n'étant pas unique, nous avons implémenté l'algorithme à partir de différents couples  $(m, g)$  que nous énonçons ci-dessous (cette liste n'est pas exhaustive).

#### Option n°1 : Lois exponentielles translatées

$$\forall (x, y) \in \mathbb{R}^2, \psi(x, y) \leq \left( \frac{10 + \sqrt{2}}{8} e^{\pi+2} \right) \times \left( 2e^{-2(x+\frac{\pi}{2})} \mathbf{1}_{x \geq -\frac{\pi}{2}} \right) \times \left( 2e^{-2(y+1)} \mathbf{1}_{y \geq -1} \right)$$

Alors, on peut prendre  $g_1$  la densité d'un vecteur aléatoire de loi  $\mathcal{TE}(2, -\frac{\pi}{2}) \otimes \mathcal{TE}(2, -1)$  et une constante  $m_1 := \frac{10+\sqrt{2}}{8} e^{\pi+2} \simeq 243,96$ .

Le ratio d'acceptation associé est :

$$\rho(x, y) = \frac{\psi(x, y)}{mg(x, y)} = \frac{\left[ \left| \sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right) \right| + 4 \cos(x)^2 + y^4 \right] \times e^{-2(|y|-y)} \mathbf{1}_{x \leq \frac{\pi}{2}} \mathbf{1}_{y \leq 1}}{\frac{10+\sqrt{2}}{2}}$$

## Option n°2 : Loi exponentielle translatée et loi de Laplace

$$\forall (x, y) \in \mathbb{R}^2, \psi(x, y) \leq \left( \frac{10 + \sqrt{2}}{4} \times e^\pi \right) \times \left( 2e^{-2(x+\frac{\pi}{2})} \mathbf{1}_{x \geq -\frac{\pi}{2}} \right) \times (e^{-2|y|})$$

Alors, on peut prendre  $g_2$  la densité d'un vecteur aléatoire de loi  $\mathcal{TE}(2, -\frac{\pi}{2}) \otimes \text{Laplace}(0, \frac{1}{2})$  et une constante  $m_2 := \left( \frac{10+\sqrt{2}}{4} \times e^\pi \right) \simeq 66,03$

Le ratio d'acceptation associé est :

$$\rho(x, y) = \frac{\psi(x, y)}{mg(x, y)} = \frac{[|\sin(\frac{2}{\pi}x^2 - \frac{\pi}{4})| + 4 \cos(x)^2 + y^4] \mathbf{1}_{x \leq \frac{\pi}{2}} \mathbf{1}_{|y| \leq 1}}{\frac{10+\sqrt{2}}{2}}$$

Le générateur de la loi de Laplace n'étant pas autorisé dans cet exercice, il nous a fallu le coder. Pour cela, on utilise la méthode de la fonction inverse. La fonction de répartition de la loi de Laplace de paramètres 0 et 1/2 est définie sur  $\mathbb{R}$

$$F(x) = 0,5[1 + \text{sgn}(x)(1 - \exp(-2|x|))]$$

Donc sa fonction réciproque définie sur  $[0, 1]$  est

$$F^{-1}(p) = \frac{1}{2} \text{sgn}(p - 0,5) \ln(1 - 2|p - 0,5|)$$

Si  $U \sim \mathcal{U}([0, 1])$  alors  $F^{-1}(U) \sim \text{Laplace}(0, \frac{1}{2})$  donc par translation, si  $U \sim \mathcal{U}([-\frac{1}{2}, \frac{1}{2}])$  alors  $F^{-1}(U) = -\frac{1}{2} \text{sgn}(U) \ln(1 - 2|U|) \sim \text{Laplace}(0, \frac{1}{2})$

## Option n°3 : Loi de Laplace et loi exponentielle tronquées

$$\forall (x, y) \in \mathbb{R}^2, \psi(x, y) \leq \left( \frac{10 + \sqrt{2}}{2} \sinh(\pi)(1 - e^{-2}) \right) \times \left( \frac{1}{\sinh(\pi)} e^{-2(x)} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \right) \times \left( \frac{1}{(1 - e^{-2})} e^{-2|y|} \mathbf{1}_{|y| \leq 1} \right)$$

Alors, on peut prendre  $g_3$  la densité d'un vecteur aléatoire de loi  $\mathcal{TE}(2, -\frac{\pi}{2}) \otimes \text{Laplace}(0, \frac{1}{2})$  de supports respectifs  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ , et  $[-1, 1]$  et une constante  $m_3 := \left( \frac{10+\sqrt{2}}{2} \right) \sinh(\pi)(1 - e^{-2}) \simeq 56,99$ .

Le ratio d'acceptation associé est :

$$\rho(x, y) = \frac{\psi(x, y)}{mg(x, y)} = \frac{[|\sin(\frac{2}{\pi}x^2 - \frac{\pi}{4})| + 4 \cos(x)^2 + y^4]}{\frac{10+\sqrt{2}}{2}}$$

Ces lois n'étant pas présentes de façon générique sur  $\mathbb{R}$ , il nous a fallu le coder. Pour cela, on utilise la méthode de la fonction inverse comme précédemment. La fonction fonction réciproque de la loi de Laplace concentrée définie sur  $[0, 1]$  est

$$F^{-1}(p) = 0.5[\log(2(1 - e^{-2})p + e^{-2})\mathbf{1}_{p < 0.5} + \mathbf{1}_{p \geq 0.5} * \log(1 + 2(1 - e^{-2})p - (1 - e^{-2}))]$$

Tandis que la fonction réciproque de la loi exponentielle concentrée définie sur  $[0, 1]$  est

$$F^{-1}(p) = \frac{1}{2} \ln(e^\pi - 2 \sinh(\pi)p)$$

#### Option n°4 : Loi uniforme et loi de Laplace

$$\forall (x, y) \in \mathbb{R}^2, \psi(x, y) \leq \left( \frac{10 + \sqrt{2}}{2} \right) \pi e^\pi \times \left( \frac{1}{\pi} \mathbf{1}_{x \in [-\frac{\pi}{2}, \frac{\pi}{2}]} \right) \times (e^{-2|y|})$$

Alors, on peut prendre  $g_4$  la densité d'un vecteur aléatoire de loi  $\mathcal{U} \left( \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \right) \otimes \text{Laplace} \left( 0, \frac{1}{2} \right)$  et une constante  $m_4 := \left( \frac{10 + \sqrt{2}}{2} \right) \pi e^\pi \simeq 414,9$

Le ratio d'acceptation associé est :

$$\rho(x, y) = \frac{\psi(x, y)}{mg(x, y)} = \frac{\left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 + y^4 \right] \times e^{-2x} \times \mathbf{1}_{y \in [-1, 1]}}{\left( \frac{10 + \sqrt{2}}{2} \right) e^\pi}$$

Nous aurions également pu ajouter une cinquième option ici où  $g$  est la densité d'un vecteur aléatoire de loi  $\mathcal{U} \left( \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \right) \otimes \mathcal{U}([-1, 1])$  mais nous ne la considérerons pas ici en raison de la pertinence de cette solution (la constante considérée était trop importante). On peut comparer ces différentes méthodes d'acceptation-rejet basées sur les densités instrumentales  $g_1, \dots, g_4$  au travers de leur bornes respectives  $m_1, \dots, m_4$ . La méthode la plus optimale est celle pour laquelle la borne  $m$  est la plus petite, soit la 3ème option ici.

**Afin de ne pas alourdir le code et le rapport, nous considererons  $m = m_3$  et  $g = g_3$  dans la suite.**

#### 3. Option n°1 : Loïs exponentielles translatées

Cette première option compte  $n_t = 169900$  simulations suivant  $g_1$  et  $n_l = 3$  passages dans la boucle while en moyenne.

#### Option n°2 : Loi exponentielle translatée et loi de Laplace

Cette option compte  $n_t = 44208$  simulations suivant  $g_2$  et  $n_l = 4$  passages dans la boucle while en moyenne.

#### Option n°3 : Loi de Laplace et loi exponentielle tronquées

Cette option compte  $n_t = 39829$  simulations suivant  $g_3$  et  $n_l = 3$  passages dans la boucle while en moyenne.

#### Option n°4 : Loi uniforme et loi de Laplace

Cette dernière option compte  $n_t = 285743$  simulations suivant  $g_4$  et  $n_l = 2$  passages dans la boucle while en moyenne. On remarque ici que cette option nécessite 7 fois plus de simulations suivant  $g_4$  que l'option n°3 ce qui correspond au rapport  $\frac{m_4}{m_3}$ . Cela confirme notre choix de l'option 3 comme étant l'option la plus optimale.

### 1.1 Méthode n°1 - Estimation de $a$

4. a) D'une part,

$$\forall (x, y) \in \mathbb{R}^2, f(x, y) = a\psi(x, y), \quad a > 0 \Rightarrow \int_{\mathbb{R}^2} f(x, y) \, dx dy = \int_{\mathbb{R}^2} a\psi(x, y) \, dx dy$$

Or,  $f$  est une densité de probabilité donc cela équivaut à

$$1 = \int_{\mathbb{R}^2} a\psi(x, y) \, dx dy$$

Par ailleurs, par définition, on a

$$\forall (x, y) \in \mathbb{R}^2, m \times g(x, y) \times \rho(x, y) = \psi(x, y)$$

Donc, par intégration,

$$\int_{\mathbb{R}^2} m \times g(x, y) \times \rho(x, y) \, dx dy = \int_{\mathbb{R}^2} \psi(x, y) \, dx dy \neq 0$$

Ainsi,

$$a = \frac{1}{\int_{\mathbb{R}^2} \psi(x, y) \, dx dy} = \frac{1}{\int_{\mathbb{R}^2} m \times g(x, y) \times \rho(x, y) \, dx dy} = \frac{1}{m \times \mathbb{E}_g [\rho(X, Y)]}$$

où  $(X, Y)$  est un vecteur aléatoire de densité  $g$ .

On peut donc estimer  $\mathbb{E}_g [\rho(X, Y)]$  à l'aide de l'estimateur de Monte-Carlo classique :

$$\bar{\rho}_n = \frac{1}{n} \sum_{i=1}^n \rho(X_k, Y_k)$$

où les vecteurs aléatoires  $(X_k, Y_k)$  sont i.i.d et de densité  $g$ . Il est possible de montrer que l'estimateur de  $a$ ,

$$\hat{b}_n := \frac{1}{m \bar{\rho}_n}$$

est biaisé. En effet, la fonction  $f(x) = 1/x$  est strictement convexe sur  $\mathbb{R}_+^*$ , donc en utilisant l'inégalité de Jensen,

$$\mathbb{E}_g [\hat{b}_n] = \frac{n}{m} \mathbb{E}_g \left[ \frac{1}{\sum_{i=1}^n \rho(X_k, Y_k)} \right] > \frac{n}{m \mathbb{E}_g [\sum_{i=1}^n \rho(X_k, Y_k)]} = a$$

De plus, les vecteurs aléatoires  $(X_k, Y_k)$  étant i.i.d et de même loi  $g$  et sachant que  $\rho(X_1, Y_1)$  est intégrable sous  $g$  (car bornée par 1), on a, par la loi forte des grands nombres,

$$\bar{\rho}_n \xrightarrow[n \rightarrow +\infty]{} \rho(X, Y) \quad \text{p.s.}$$

donc, par continuité de la fonction  $g(x) = \frac{1}{mx}$  sur  $\mathbb{R}_+^*$  on obtient

$$\hat{b}_n \xrightarrow[n \rightarrow +\infty]{} a \quad \text{p.s.}$$

L'estimateur  $\hat{b}_n$  est donc consistant.

Les variables aléatoires  $\rho(X_k, Y_k)$  sont i.i.d et de carré intégrable (car elles sont bornées par 1) donc en utilisant le Théorème Central limite,

$$\sqrt{n} (\bar{\rho}_n - \mathbb{E}_g [\rho(X, Y)]) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, \sigma^2) \quad \text{où } \sigma^2 = \mathbb{V}ar_g[\rho(X, Y)]$$

D'autre part, la fonction  $g(x) = \frac{1}{mx}$  est de classe  $\mathcal{C}^1$  sur  $\mathbb{R}_+^*$  de dérivée  $g'(x) = -\frac{1}{mx^2}$  définie sur  $\mathbb{R}_+^*$ . En utilisant la méthode delta, on a donc

$$\sqrt{n} (\hat{b}_n - a) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, \sigma^2 (a^2 m)^2)$$

On introduit un estimateur sans biais de  $\sigma^2$  :  $\hat{\sigma}_{n-1}^2 := \frac{1}{n-1} \sum_{i=1}^n (\rho(X_k, Y_k) - \hat{\rho}_n)^2$

Or, le théorème de continuité permet d'écrire :

$$\sqrt{\frac{\sigma^2 a^4}{\hat{\sigma}_{n-1}^2 \hat{b}_n^4}} \xrightarrow[n \rightarrow +\infty]{loi} 1$$

Le théorème de Slutsky permet alors de conclure

$$\frac{1}{\hat{b}_n^2 m} \sqrt{\frac{n}{\hat{\sigma}_{n-1}^2}} (\hat{b}_n - a) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, 1)$$

Si on note  $q_{1-\frac{\alpha}{2}}$  le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi normale  $\mathcal{N}(0, 1)$  l'intervalle de confiance asymptotique de niveau  $1 - \alpha$  est

$$IC_{1-\alpha} = \left[ \hat{b}_n - \left( \hat{b}_n^2 m \right) \sqrt{\frac{\hat{\sigma}_{n-1}^2}{n}} q_{1-\frac{\alpha}{2}}, \hat{b}_n + \left( \hat{b}_n^2 m \right) \sqrt{\frac{\hat{\sigma}_{n-1}^2}{n}} q_{1-\frac{\alpha}{2}} \right]$$

b) On trouve en pratique  $\hat{b}_n = 0,07$  et un intervalle de confiance :

$$IC_{95\%} = [0.06869079, 0.06987027]$$

c) Afin de ne pas avoir à simuler des variables aléatoires supplémentaires, on utilise la méthode du bootstrap. Cette méthode utilise un ré-échantillonnage de nos données afin d'estimer ici le biais de notre estimateur  $\hat{b}_n$ . Le biais observé, est en pratique de l'ordre de  $10^{-5}$ .

5. (a) La variable aléatoire  $T$  représente le temps d'attente avant le premier succès dans une suite d'épreuves de Bernoulli indépendantes. C'est donc une loi géométrique de paramètre  $\frac{1}{M}$ . On en déduit en particulier que  $\mathbb{E}[T] = M = am$  ou encore  $a = \frac{\mathbb{E}[T]}{m}$ .

A partir de l'estimateur de Monte-Carlo classique de  $\mathbb{E}[T]$ , on définit un autre estimateur de  $a$

$$\hat{a}_n := \frac{1}{m} \left( \frac{1}{n} \sum_{k=1}^n T_k \right) \text{ où } \forall k \in \{1, \dots, n\}, \quad T_k \sim \mathcal{G} \left( \frac{1}{M} \right)$$

Les variables aléatoires  $(T_n)_{n \in \mathbb{N}}$  représentent le temps d'attente avant le  $n^{me}$  succès. Notons alors,  $\bar{T}_n := \frac{1}{n} \sum_{k=1}^n T_k$

Tout d'abord, on remarque que  $\hat{a}_n$  est sans biais. En effet,

$$\mathbb{E}[\hat{a}_n] = \mathbb{E} \left[ \frac{1}{mn} \sum_{k=1}^n T_k \right] = \frac{1}{m} \mathbb{E}[T_1] = a$$

De plus, les variables aléatoires  $(T_n)_{n \in \mathbb{N}}$  sont i.i.d et de carré intégrables donc par la Loi Forte des Grands Nombres,

$$\bar{T}_n \xrightarrow[n \rightarrow +\infty]{} a \quad \text{p.s}$$

On en déduit par le Théorème Central Limite,

$$\sqrt{n} (\bar{T}_n - am) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, am(am - 1))$$

La méthode delta appliquée à la fonction dérivable  $h(x) := \frac{x}{m}$  de dérivée non nulle  $h'(x) = \frac{1}{m}$ , donne

$$\sqrt{n}(\hat{a}_n - a) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}\left(0, a(am - 1) \times \frac{1}{m}\right)$$

D'où, par continuité de  $x \rightarrow \frac{1}{\sqrt{x(am-1)}}$  sur  $]1, +\infty[$  et par le théorème de Slutsky, on en déduit que

$$\sqrt{\frac{nm}{\hat{a}_n(\hat{a}_nm - 1)}}(\hat{a}_n - a) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, 1)$$

En notant  $q_{1-\frac{\alpha}{2}}$  le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi normale  $\mathcal{N}(0, 1)$  l'intervalle de confiance asymptotique de niveau  $1 - \alpha$  est

$$IC_{1-\alpha} = \left[ \hat{a}_n - \sqrt{\frac{\hat{a}_n(\hat{a}_nm - 1)}{nm}} q_{1-\frac{\alpha}{2}}, \hat{a}_n + \sqrt{\frac{\hat{a}_n(\hat{a}_nm - 1)}{nm}} q_{1-\frac{\alpha}{2}} \right]$$

(b) On trouve en pratique  $\hat{a}_n = 0,07$  et un intervalle de confiance :

$$IC_{95\%} = [0.06869028, 0.07106066]$$

6. On s'intéresse au rapport des coûts pour lesquels  $\hat{b}_n$  et  $\hat{a}_n$  atteignent la même précision. Pour cela, il nous faut calculer la variance de ces estimateurs. Tout d'abord,

$$\mathbb{V}ar[\hat{a}_n] = \frac{1}{m^2 n} \mathbb{V}ar[T] \simeq 3.66 \times 10^{-7}$$

D'autre part, pour  $n$  assez grand (on considère  $n \geq 30$ ),  $\frac{1}{\hat{b}_n^2 m} \sqrt{\frac{n}{\hat{\sigma}_{n-1}^2}} (\hat{b}_n - a)$  se comporte comme une loi  $\mathcal{N}(0, 1)$  donc

$$\mathbb{V}ar \left[ \frac{1}{\hat{b}_n^2 m} \sqrt{\frac{n}{\hat{\sigma}_{n-1}^2}} (\hat{b}_n - a) \right] = 1 \Leftrightarrow \mathbb{V}ar[\hat{b}_n] = (\hat{b}_n m)^2 \times \frac{\hat{\sigma}_{n-1}}{n} \simeq 9.05 \times 10^{-8}$$

Le rapport des variances des 2 estimateurs nous indique que l'estimateur  $\hat{b}_n$  est 25 fois plus efficace que l'estimateur  $\hat{a}_n$  en terme de variance. En estimant le rapport des coûts, on obtient que l'estimateur  $\hat{b}_n$  est environ 16 fois plus efficace que l'estimateur  $\hat{a}_n$  en terme d'efficacité relative.

Par ailleurs, si  $\epsilon = 0.01$  alors le nombre de tirages suivant  $f$  et  $g$  nécessaire pour obtenir une précision à  $\epsilon$  près pour un intervalle de confiance de  $\hat{a}_n$  et  $\hat{b}_n$  à 95% est respectivement de  $n_f := 424$  et  $n_g := 1$ .

7. (a) En effectuant des intégrations par parties successives, on remarque que

$$\begin{aligned} \int_{\mathbb{R}} \psi(x, y) dy &= \int_{\mathbb{R}} \left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 + y^4 \right] e^{-2(x+|y|)} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \mathbf{1}_{|y| \leq 1} dy \\ &= \left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 \right] e^{-2x} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \left[ \int_{\mathbb{R}} e^{-2|y|} \mathbf{1}_{|y| \leq 1} dy + \int_{\mathbb{R}} y^4 e^{-2|y|} \mathbf{1}_{|y| \leq 1} dy \right] \\ &= \left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 \right] e^{-2x} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \left[ 2 \int_0^1 e^{-2y} dy + 2 \int_0^1 y^4 e^{-2y} dy \right] \\ &= \left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 \right] e^{-2x} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \left[ (1 - e^{-2}) + \left( \frac{3 - 21e^{-2}}{2} \right) \right] \end{aligned}$$

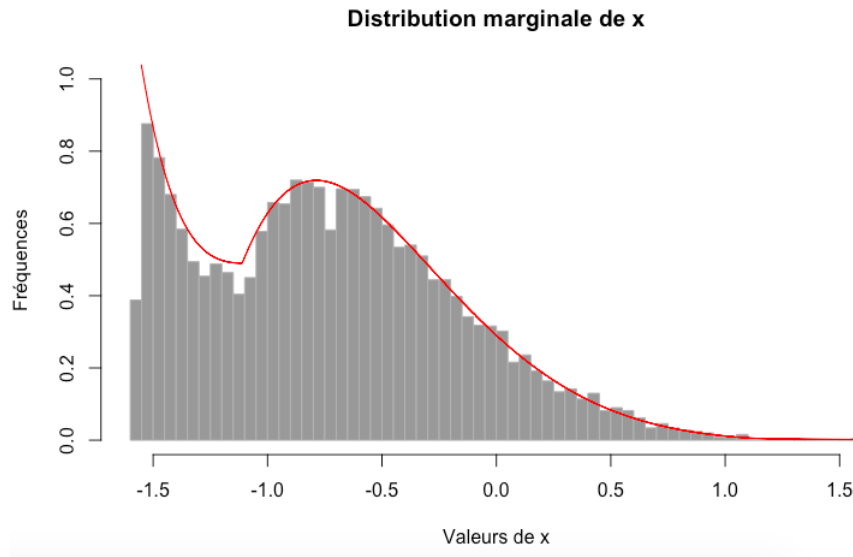


On s'intéresse alors à l'estimateur  $\hat{f}_{X,n}(x)$  de  $f_X(x)$  défini ci-dessous à partir de  $\hat{a}_n$  défini en question 5 :

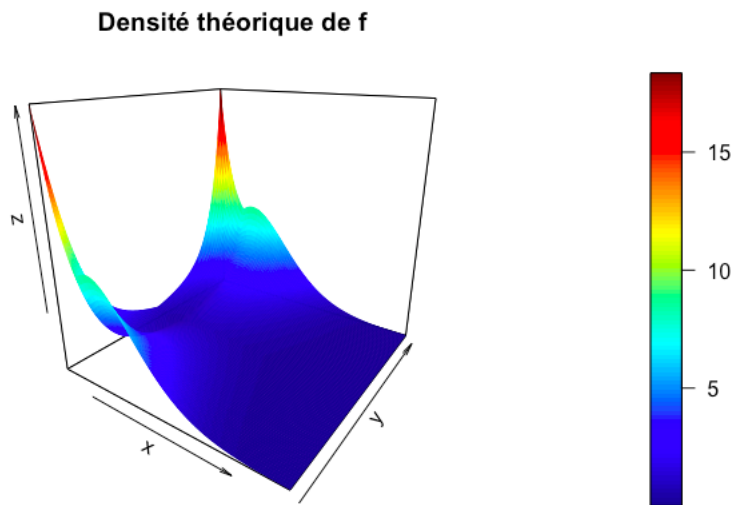
$$\forall x \in \mathbb{R} \quad \hat{f}_{X,n}(x) := \hat{a}_n \times \left( \left[ \left| \sin \left( \frac{2}{\pi} x^2 - \frac{\pi}{4} \right) \right| + 4 \cos(x)^2 \right] e^{-2x} \mathbf{1}_{|x| \leq \frac{\pi}{2}} \left[ (1 - e^{-2}) + \left( \frac{3 - 21e^{-2}}{2} \right) \right] \right)$$

Cet estimateur est bien consistant par consistance de  $\hat{a}_n$ .

- (b) A partir de la densité marginale calculée dans la question précédente, on compare la distribution marginale de l'échantillon  $z$  représentée sous la forme d'un histogramme à l'estimateur  $\hat{f}_{X,n}(x)$  représenté en rouge.



A l'aide de l'estimateur  $\hat{a}_n$ , on représente la densité théorique de  $f$  ci-dessous.



## 1.2 Méthode n°2 - Estimateur ponctuel

8. On note  $\hat{W}^k(X_k, Y_k) := \frac{\psi(x, Y_k)\omega(X_k)}{\psi(X_k, Y_k)}$  pour  $k \in \{1, \dots, n\}$ . Cet estimateur est intégrable. En effet,

$$\begin{aligned}\mathbb{E}[\hat{W}^k(X_k, Y_k)] &= \mathbb{E}[\hat{W}^k(X_k, Y_k)] = \iint \frac{\psi(x, y)\omega(z)}{\psi(z, y)} f_{X,Y}(z, y) \, dz \, dy \\ &= \iint \frac{a\psi(x, y)\omega(z)\psi(y, z)}{\psi(z, y)} \, dz \, dy \\ &= \int a\psi(x, y) \, dy \int \omega(z) \, dz \\ &= \int f_{X,Y}(x, y) \, dy \\ &= f_X(x)\end{aligned}$$

Ainsi, par la loi forte des grands nombres  $\hat{\omega}_n(x) := \overline{\hat{W}^k(X_k, Y_k)} \xrightarrow[n \rightarrow +\infty]{p.s} f_X(x)$ .

En prenant,  $\omega$  tq  $\hat{W}^k$  soit de carré intégrable, on peut appliquer le TCL, pour  $x \in \mathbb{R}$

$$\sqrt{n}(\hat{\omega}_n(x) - f_X(x)) \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, \text{Var}(\hat{W}^k))$$

On note  $\hat{\sigma}_1^2 := \frac{1}{n-1} \sum_{k=1}^n (\omega_1 - \hat{\omega}_n(x))^2$  estimateur non biaisé de  $\text{Var}(\omega_1)$  et p.s convergent vers  $\text{Var}(W^k)$  (par la LGN).

Ainsi, par le théorème de Slutsky, et continuité de  $(x, y) \mapsto \frac{x}{y}$ , on a

$$\sqrt{n} \frac{(\hat{\omega}_n(x) - f_X(x))}{\hat{\sigma}_1^2} \xrightarrow[n \rightarrow +\infty]{loi} \mathcal{N}(0, 1).$$

En notant  $q_{1-\frac{\alpha}{2}}$  le quantile d'ordre  $1 - \frac{\alpha}{2}$  d'une  $\mathcal{N}(0, 1)$ , on retrouve un IC asymptotique bilatéral symétrique pour  $f_X(x)$  :

$$f_X(x) \in \left[ \hat{\omega}_n(x) - \sqrt{\frac{\hat{\sigma}_1^2}{n}} q_{1-\frac{\alpha}{2}}; \hat{\omega}_n(x) + \sqrt{\frac{\hat{\sigma}_1^2}{n}} q_{1-\frac{\alpha}{2}} \right]$$

9. Calculons la variance de l'estimateur pour trouver un critère de minimisation .

$$\text{var}(\hat{\omega}_n(x)) = \frac{1}{n} \text{var}(\hat{W}^1) \tag{1}$$

$$= \frac{1}{n} [\mathbb{E}[(\hat{W}^1)^2] - f_X(x)^2] \tag{2}$$

$$= \frac{1}{n} \iint \frac{(a\psi(x, y)\omega(z) - af_X(x)\psi(z, y))^2}{a\psi(z, y)} \, dz \, dy \tag{3}$$

L'intégrande étant toujours positive, il faut et suffit que celle-ci soit nulle pp.(z,y) pour que la variance soit nulle. On cherche donc  $\omega$  densité tel que pour (presque tout)  $z \in [-\frac{\pi}{2}; \frac{\pi}{2}]$  :

$$a\psi(x, y)\omega(z) - af_X(x)\psi(z, y) = 0$$

$$\text{i.e } \omega(z) = \frac{f_X(x)f_{X,Y}(z, y)}{f_{X,Y}(x, y)} \quad \text{car } f_{X,Y}(x, y) \neq 0 \text{ pour tout } (x, y)$$

Dans le cas où  $X$  et  $Y$  sont indépendants,  $\omega$  est bien une densité (d'intégrale 1). Dans ce même cas, on a alors  $\omega(z) = \frac{f_X(x)f_X(z)f_Y(y)}{f_X(x)f_Y(y)} = f_X(z)$ . En pratique, ce choix n'est pas exploitable, car dépend de la quantité à estimer. Cependant, cela donne une idée sur le choix de  $\omega$  : on le choisit tel que  $\frac{\omega(z)}{f_X(z)} \approx \text{constante}$  et de variance finie.

10. En prenant  $\omega(x) := \frac{1}{\sinh(\pi)} e^{-2x} \mathbf{1}_{|x| \leq \frac{\pi}{2}}$ , on obtient un rapport de  $\frac{\omega(z)}{f_X(z)} = (|\sin(\frac{2}{\pi}z^2 - \frac{\pi}{4})| + 4 \times \cos(z)^2) \times c \times \sinh(\pi) \times \hat{a}_n$  pour  $z \in [-\frac{\pi}{2}; \frac{\pi}{2}]$ , et  $c = 1 - e^{-2} + 3 - 21\frac{e^{-2}}{2}$ . Ce rapport est compris entre 0.5 et 3.5, on peut donc le considérer comme étant à peu près constant. On obtient alors  $\hat{\omega}_n(-1) \approx 0.60$  et un  $IC_{95\%}$  de  $[0.594, 0.612]$ .
11. On s'intéresse au rapport des coûts pour lesquels  $\hat{\omega}_n(-1)$  et  $\hat{f}_{X,n}(x)$  atteignent la même erreur quadratique. Si  $mse = 2.85e - 05$ , alors le nombre de tirages suivant  $\hat{f}_X$  et  $\hat{W}$  nécessaire est  $n_{\text{estimateur}} = \frac{\text{var}(\text{estimateur})}{mse}$ , soit  $n_{f_X} := 10000$  et  $n_{\hat{W}} := 7583$ .  
Par ailleurs, pour obtenir  $n = 10000$  simulations suivant  $f_X$ , il nous faut 10000 simulations de  $\hat{a}_n$ , donc le coût de  $\hat{f}_X$  est  $C_1 := C_{\hat{a}_1} = \frac{n_t \times n_l}{n} \approx 13.262$ .  
Le coût de simulation de  $\hat{W}$  est  $C_2 := \frac{n_t \times n_l}{n} \approx 13.262$ . Ainsi, le rapport des coûts entre les 2 méthodes est de  $R = \frac{C_2 n_{f_X}}{C_1 n_{\hat{W}}} = 1.31$ . La méthode par estimation ponctuelle est donc légèrement plus efficace que celle de Monte-Carlo classique.

## 2 Exercice 2

1. L'objectif de cet exercice est d'estimer la valeur de

$$\delta = \mathbb{E}[\min(3, \frac{1}{3} \sum_{k=1}^3 e^{-X_k})]$$

où  $\mathbf{X} = (X_1, X_2, X_3)$  est distribué suivant la loi  $\mathcal{N}(\mu, \Sigma)$ .

La première problématique est d'abord la simulation de ces variables aléatoires. Pour générer un vecteur suivant cette loi, on s'appuie sur la proposition 2.2 du cours : la matrice de covariance  $\Sigma$  étant symétrique définie positive (ses valeurs propres sont toutes strictement positives), on note  $L$  une matrice triangulaire inférieure tel que  $\Sigma = LL^t$ . On simule  $Z$  un vecteur aléatoire de loi  $\mathcal{N}(\mathbf{0}_d, \mathbf{I}_d)$ , alors le vecteur  $\mathbf{X} = \mu + L\Sigma$  est une réalisation suivant la loi normale multivariée  $\mathcal{N}(\mu, \Sigma)$ .

2. (a)  $\delta$  étant déjà donné sous la forme d'une espérance, on définit

$$h : \begin{cases} \mathbb{R}^3 & \longrightarrow \mathbb{R} \\ (X_1, X_2, X_3) & \longmapsto \min(3, \frac{1}{3} \sum_{k=1}^3 e^{-X_k}) \end{cases}$$

où  $\mathbf{X} = (X_1, X_2, X_3)$  suit la loi  $\mathcal{N}(\mu, \Sigma)$ .

L'estimateur de Monte-Carlo s'écrit alors  $\bar{\delta}_n = \frac{1}{n} \sum_{i=1}^n h(X_1^i, X_2^i, X_3^i)$  où les  $\mathbf{X}^i$  suivent la même loi que  $\mathbf{X}$ .

- (b) L'erreur quadratique moyenne associée à cet estimateur est de  $1.68 * 10^{-6}$

3. (a) On pose

$$A : \begin{cases} \mathbb{R}^3 & \longrightarrow \mathbb{R} \\ X & \longmapsto 2\mu - X \end{cases}$$

Alors pour  $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$ ,  $A(\mathbf{X}) \sim \mathcal{N}(\mu, \Sigma)$ . En effet, il s'agit d'une combinaison linéaire du vecteur gaussien  $(\mu, \mathbf{X})$  de moyenne  $2\mu - \mu = \mu$  et de matrice de variance  $\mathbf{0}_d + \Sigma = \Sigma$ .

Par ailleurs, la transformation  $A$  est décroissante tandis que  $h$  est monotone (décroissante non constante) en chacune de ses coordonnées, donc  $Cov[h(X), A(h(X))] < 0$  et ainsi  $Var[\hat{\delta}_n] \leq Var[\bar{\delta}_n] / 2$ .

- (b) L'erreur quadratique moyenne associée à cet estimateur est de l'ordre de  $3.4e-8$  et  $R_1 \approx 48.25$ . La réduction est significative : à coûts de calcul équivalents et même nombre de simulations, l'estimateur  $\bar{\delta}_n$  est 50 fois plus efficace que  $\hat{\delta}_n$ .

4. En utilisant le fait que le vecteur  $\mathbf{X}$  est centré autour de  $\mathbf{0}_3$ , on effectue un DL de l'exponentielle autour de 0 à l'ordre 2 :

$$e^{-X_k} = 1 + X_k + \frac{X_k^2}{2} + o(X_k^3) \quad \text{pour } k \in \{1, 2, 3\}$$

On pose alors :

$$h_0(\mathbf{X}) = \frac{1}{3} \sum_{i=1}^3 1 + X_k + \frac{X_k^2}{2}$$

Alors,

$$\begin{aligned}
\mathbb{E}[h_0(\mathbf{X})] &= \frac{1}{3} \sum_{i=1}^3 \mathbb{E} \left[ 1 + X_k + \frac{X_k^2}{2} \right] \\
&= \frac{1}{3} \sum_{i=1}^3 1 + \mathbb{E}[X_k] + \mathbb{E} \left[ \frac{X_k^2}{2} \right] \\
&= \frac{2.8805}{3}
\end{aligned}$$

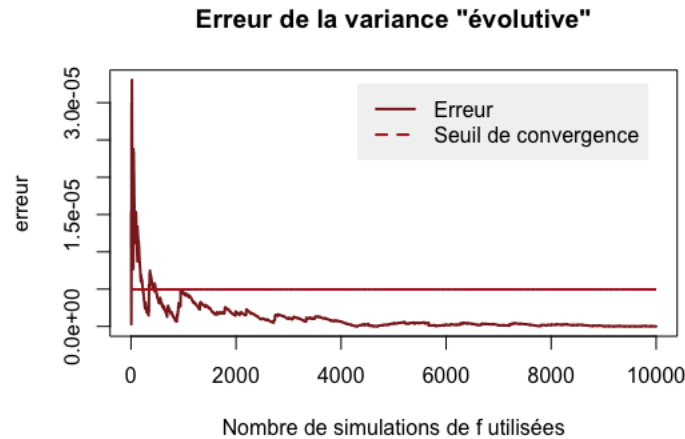
Par ailleurs,  $h_0$  est bien de carré intégrable. En effet,

$$\begin{aligned}
\mathbb{E}[(h_0(\mathbf{X}^2))] &= \frac{1}{9} \mathbb{E} \left[ \sum_{k=1}^3 \left( 1 + X_k + \frac{X_k^2}{2} \right)^2 \right] \\
&= \frac{1}{9} \sum_{k=1}^3 \sum_{j=1}^3 \mathbb{E} \left[ \left( 1 + X_k + \frac{X_k^2}{2} \right) \left( 1 + X_j + \frac{X_j^2}{2} \right) \right] \\
&\leq \frac{1}{9} \sum_{k=1}^3 \sum_{j=1}^3 \mathbb{E} \left[ \left( 1 + X_k + \frac{X_k^2}{2} \right)^2 \right] \mathbb{E} \left[ \left( 1 + X_j + \frac{X_j^2}{2} \right)^2 \right]
\end{aligned}$$

où  $\mathbb{E} \left[ \left( 1 + X_k + \frac{X_k^2}{2} \right)^2 \right]$  est une fonction linéaire des moments d'ordre 1, 2, 3 et 4 d'une loi normale, donc bien définie.  $h_0$  ainsi défini, on trouve une corrélation  $\approx 0.99$ . Ainsi, l'estimateur par la méthode de la variable de contrôle simple est donné par :

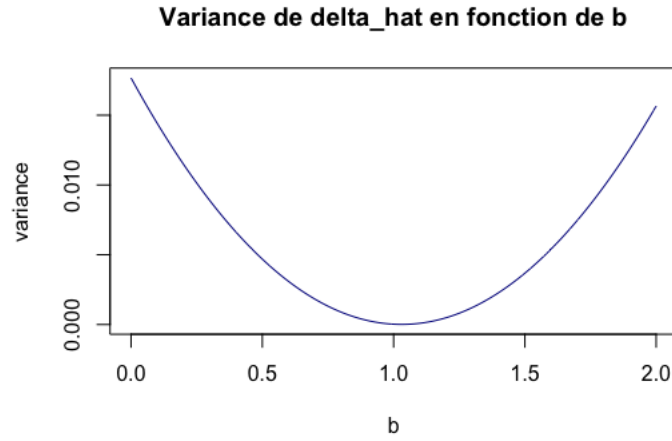
$$\hat{\delta}_n(b) = \frac{1}{n} \sum_{k=1}^n (\mathbf{X}_k) - b * (h_0(\mathbf{X}_k) - \mathbb{E}[h_0(\mathbf{X}_k)])$$

5. L'idée ici est d'utiliser la méthode burn-in pour évaluer le  $b_{opt}$  qui minimisera au plus la variance, puis  $\hat{\delta}_n(b_{opt})$  pour éviter l'apparition d'un biais dans l'estimation. Afin de trancher  $\mathbf{z}$  au bon endroit, On compare la variance évolutive de  $\hat{\delta}_n(b)$  en fonction de  $n$  avec celle de l'estimateur final. On coupera au premier indice à partir duquel la différence entre les 2 valeurs est plus petite qu'un certain epsilon (que l'on fixe à  $5 \times 10^{-6}$  dans le code).



On détermine graphiquement  $nbsimu \approx 300$ .

On estime ensuite graphiquement la valeur de  $b$  pour laquelle la variance de  $\hat{\delta}_{n-nbsimu}(b)$  est minimale.



On trouve alors  $b_{opt} = 1.03$ ,  $\hat{\delta}_n(b_{opt}) \approx 0.95$  et  $mse(\hat{\delta}_n(b_{opt})) \approx 8.75 \times 10^{-10}$ . Cet estimateur fait donc 40 fois mieux que celui de la variable antithétique, et 1900 fois mieux que l'estimateur de Monte-Carlo classique !

### 3 Exercice 3

1. L'objectif de cet exercice est d'estimer la valeur de

$$\delta = \mathbb{E}[S] \quad \text{avec} \quad S = \sum_{i=1}^Y \log(X_i + 1)$$

où  $Y$  est distribué suivant la loi géométrique de paramètre  $p = 0.2$  et  $(X_n)_{n \geq 1}$  est une suite de variables aléatoires i.i.d suivant la loi gamma  $\Gamma(2, 2)$ .

Pour cela, la première étape est la simulation de ces variables aléatoires. La loi géométrique définie dans l'exercice que sur  $\mathbb{N}^*$  ne correspond pas à la loi géométrique définie sur  $\mathbb{R}$ . C'est pourquoi, nous avons utilisé la méthode de la fonction inverse (que l'on notera  $F^{\leftarrow}$ ) pour simuler cette variable aléatoire. D'après l'exemple 2.1 du cours, on a  $\forall u \in [0, 1]$ ,

$$\begin{aligned} F^{\leftarrow}(u) &:= \left\{ n \in \mathbb{N}^* : \sum_{i=1}^{n-1} p(1-p)^{i-1} < u \leq \sum_{i=1}^n p(1-p)^{i-1} \right\} \\ &= \{ n \in \mathbb{N}^* : 1 - (1-p)^{n-1} < u \leq 1 - (1-p)^n \} \\ &= \{ n \in \mathbb{N}^* : (1-p)^n \leq 1-u < (1-p)^{n-1} \} \\ &= \{ n \in \mathbb{N}^* : n \ln(1-p) \leq \ln(1-u) < (n-1) \ln(1-p) \} \\ &= \left\{ n \in \mathbb{N}^* : n-1 < \frac{\ln(1-u)}{\ln(1-p)} \leq n \right\} \\ &= \left\lceil \frac{\ln(1-u)}{\ln(1-p)} \right\rceil \end{aligned}$$

Alors, d'après le lemme 2.1 du cours, si  $U \sim \mathcal{U}([0, 1])$  alors  $F^{\leftarrow}(U) \sim \mathcal{G}(p)$ .

On définit l'estimateur de Monte-Carlo classique tel que :

$$\bar{\delta}_n := \frac{1}{n} \sum_{i=1}^n S_i \quad \text{où les } S_i \text{ suivent la même loi que } S$$

L'erreur quadratique moyenne associée à cet estimateur est :

$$\mathbb{V}ar[\bar{\delta}_n] = \frac{1}{n} \mathbb{V}ar[S] \simeq 8,5 \times 10^{-4}$$

2. (a) On choisit

- $Y$  comme variable de stratification
- On partitionne  $[0, 1]$  de la façon suivante :

$$\forall k \in \{1, \dots, L-1\} \quad D_k = \{k\} = F^{\leftarrow}([1 - (1-p)^{k-1}, 1 - (1-p)^k])$$

$$D_L = F^{\leftarrow}([1 - (1-p)^{L-1}, 1])$$

de sorte que  $\mathcal{D} = (D_1, \dots, D_L)$  forme une partition de  $\mathbb{N}^*$ .

On a alors

$$\begin{aligned}\forall k \in \{1, \dots, L-1\} \quad p_k &:= \mathbb{P}(Y \in (D_k)) = \mathbb{P}(Y = k) = p(1-p)^{k-1} \\ p_L &:= \mathbb{P}(Y \in (D_L)) = \mathbb{P}(Y \geq L) = (1-p)^{L-1}\end{aligned}$$

Par ailleurs, on choisit l'allocation proportionnelle des  $n$  tirages c'est à dire que pour chaque strate  $D_k$ , on prend

$$\begin{aligned}\forall k \in \{1, \dots, L-1\} \quad n_k &= n\mathbb{P}(Y \in (D_k)) = np(1-p)^{k-1} \\ n_L &= n\mathbb{P}(Y \in (D_L)) = n(1-p)^{L-1}\end{aligned}$$

Puis, pour chaque strate  $D_k$ , on considère  $(S_n^{(k)})_{n \in \mathbb{N}}$  une suite de variables aléatoires i.i.d suivant la loi conditionnelle  $\mathcal{L}(S|Y \in (D_k))$ . Pour  $k \in \{1, \dots, L-1\}$ , cela correspond à la simulation d'une somme déterministe. En revanche, pour  $k = L$ , cette simulation n'est plus aussi évidente. Il s'agit, en effet, de simuler des variables aléatoires suivant la loi conditionnelle  $\mathcal{L}(S|Y \geq L)$ . L'estimateur stratifié s'écrit donc

$$\begin{aligned}\hat{\delta}_n(n_1, \dots, n_L) &:= \sum_{k=1}^L \frac{1}{n} \sum_{i=1}^{n_k} S_i^k \\ &= \sum_{k=1}^{L-1} p_k \left( \frac{1}{n_k} \sum_{i=1}^{n_k} S_i^k \right) + p_L \left( \frac{1}{n_L} \sum_{i=1}^{n_L} S_i^L \right)\end{aligned}$$

On remarque les moyennes empiriques  $\bar{S}_{n_k}^k$  pour  $k = 1, \dots, L$ .

(b) On a par indépendance des strates

$$\mathbb{V}ar[\hat{b}_n] = \sum_{k=1}^{L-1} p_k^2 \left( \frac{1}{n_k} \mathbb{V}ar[S_{n_k}^{(k)}] \right) + p_L^2 \left( \frac{1}{n_L} \mathbb{V}ar[S_{n_L}^{(L)}] \right)$$

L'erreur quadratique moyenne associée à cet estimateur est donc d'environ  $1,1 \times 10^{-4}$  soit environ 7,5 fois moins que l'estimateur de Monte-Carlo classique. Ainsi, l'estimateur stratifié est presque 8 fois plus efficace en terme de variance que l'estimateur classique.

En théorie, l'estimateur stratifié avec allocation proportionnelle est au moins aussi efficace que l'estimateur de Monte-Carlo classique. A l'aide de la librairie *microbenchmark*, nous avons calculé les coûts de calcul des 2 estimateurs. On obtient ainsi que l'estimateur stratifié est environ 10 fois plus performant que l'estimateur de Monte-Carlo classique.