

# Projet Monte Carlo

## I. Algorithme du rejet

### Question 1 :

On souhaite obtenir une réalisation suivant la loi de densité  $f$ .

On sait que  $f(x) = c \cdot \tilde{f}(x)$  et  $\tilde{f}(x) \leq M \cdot g(x)$ , d'où  $f(x) \leq c \cdot M \cdot g(x)$ , avec  $g(x)$  une densité.

Si  $c \cdot M \geq 1$ , alors on peut appliquer l'algorithme du rejet immédiatement à  $\tilde{f}$ .

Si  $c \cdot M \leq 1$ , alors en particulier,  $c \cdot M \leq M'$ , avec  $M' \geq 1$ , d'où  $f(x) \leq M' \cdot g(x)$ , donc on peut également appliquer le rejet à  $\tilde{f}$ .

### Question 2 :

On cherche les constantes  $M_1, M_2$  et la densité  $g$  telles que  $\tilde{f}_1(x) \leq M_1 \cdot g(x)$  et  $\tilde{f}_2(x) \leq M_2 \cdot g(x)$ .

On a :

$$\tilde{f}_1(x) = \exp \left\{ -\frac{1}{2} \left( \frac{x_1^2}{4} + x_2^2 \right) \right\} \mathbb{I}_{\{|x_2| \leq 1\}} \leq \exp \left\{ -\frac{1}{2} \left( \frac{x_1^2}{4} + x_2^2 \right) \right\} = \frac{4\pi}{4\pi} * \exp \left\{ -\frac{1}{2} \left( \frac{x_1^2}{4} + x_2^2 \right) \right\}$$

D'où  $M_1 = 4\pi$  et  $g$  qui est la densité d'une  $\mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \right)$ .

En effet,  $\frac{1}{2\pi\sqrt{4}\sqrt{1}} * \exp \left\{ -\frac{1}{2} \left( \frac{x_1^2}{4} + x_2^2 \right) \right\}$  est la densité d'une loi normale  $\mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \right)$ .

Par un raisonnement analogue, on a :

$$\tilde{f}_2(x) = \left\{ \cos^2(y_1) + \frac{1}{2} \sin^2(3y_2) \cos^4(y_1) \right\} \exp \left\{ -\frac{1}{2} \left( \frac{x_1^2}{4} + x_2^2 \right) \right\} \leq \frac{3}{2} * M_1 * g$$

Car  $\cos^2(y_1) + \frac{1}{2} \sin^2(3y_2) \cos^4(y_1) \leq 1 + \frac{1}{2}$ , puisque  $0 \leq \cos(x), \sin(x) \leq 1, \forall x$ .

D'où  $M_2 = \frac{3}{2} * M_1 = 6\pi$ .

### Question 3 :

Nous avons montré qu'il existe des constantes  $M_1$  et  $M_2$  positives et une densité  $g$  telles que l'on ait  $\tilde{f}_1(x) \leq M_1 \cdot g(x)$  et  $\tilde{f}_2(x) \leq M_2 \cdot g(x)$ . D'après la [question 1](#), nous pouvons simuler  $f_1$  et  $f_2$  grâce à l'algorithme du rejet.

### Question 4 :

Calculons les marginales de  $f_1$ . Pour cela, calculons tout d'abord la constante de normalisation de  $f_1$  en fonction de la fonction de répartition  $\Phi$  de la loi normale  $\mathcal{N}(0,1)$ .

On cherche donc  $C$  tel que  $\int_{\mathbb{R}^2} C \exp \left\{ -\frac{1}{2} \left( \frac{x^2}{4} + y^2 \right) \right\} \mathbb{I}_{\{|y| \leq 1\}} dx dy = 1$

$$\begin{aligned} \int_{\mathbb{R}^2} \exp \left\{ -\frac{1}{2} \left( \frac{x^2}{4} + y^2 \right) \right\} \mathbb{I}_{\{|y| \leq 1\}} dx dy &= 2\pi \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{8}} \left( \int_{-1}^1 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right) dx \\ &= 2\pi(2\Phi(1) - 1) \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{8}} dx \end{aligned}$$

On effectue le changement de variable  $x \mapsto 2x$ . Les bornes restent inchangées. On obtient :

$$\begin{aligned} &= 2\pi(2\Phi(1) - 1) \int_{\mathbb{R}} \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= 4\pi(2\Phi(1) - 1) \end{aligned}$$

D'où  $C \cdot 4\pi(2\Phi(1) - 1) = 1$ , ce qui donne la constante de normalisation  $C = \frac{1}{4\pi(2\Phi(1)-1)}$ .

Calculons maintenant les marginales de  $f_1$ . Première marginale :

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} \tilde{f}_1(x, y) dy \\ &= C \cdot \sqrt{2\pi} \int_{\mathbb{R}} e^{-\frac{x^2}{8}} \cdot \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \mathbb{I}_{\{|y| \leq 1\}} dy \\ &= C \cdot \sqrt{2\pi} \cdot e^{-\frac{x^2}{8}} \cdot (2\Phi(1) - 1) \\ &= \frac{e^{-\frac{x^2}{8}}}{2\sqrt{2\pi}} \end{aligned}$$

Donc la première marginale de  $f_1$  est une normale  $\mathcal{N}(0, 4)$ . Calculons désormais la seconde marginale de  $f_1$ .

$$\begin{aligned} f_Y(y) &= C \int_{\mathbb{R}} \tilde{f}_1(x, y) dx \\ &= C \cdot 2\sqrt{2\pi} \int_{\mathbb{R}} e^{-\frac{x^2}{8}} \cdot \frac{e^{-\frac{y^2}{2}}}{2\sqrt{2\pi}} \mathbb{I}_{\{|y| \leq 1\}} dx \\ &= C \cdot 2\sqrt{2\pi} e^{-\frac{y^2}{2}} \cdot \mathbb{I}_{\{|y| \leq 1\}} \int_{\mathbb{R}} \frac{e^{-\frac{x^2}{8}}}{2\sqrt{2\pi}} dx \\ &= 2C \cdot \sqrt{2\pi} e^{-\frac{y^2}{2}} \cdot \mathbb{I}_{\{|y| \leq 1\}} \\ &= \frac{e^{-\frac{y^2}{2}} \cdot \mathbb{I}_{\{|y| \leq 1\}}}{\sqrt{2\pi}(2\Phi(1) - 1)} \end{aligned}$$

## II. Retour sur les méthodes de réduction de variance

### Premier Cas

Notons dans ce qui suit  $X_{1,i}$  les tirages iid de  $X_1$ , et  $X_{2,i}$  les tirages iid de  $X_2$ .

#### Question 5.a :

La méthode de Monte-Carlo Classique donne par définition l'estimateur suivant :

$$p = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq 5\}}$$

#### Question 5.b :

On sait que  $-X_1 \stackrel{loi}{\cong} X_1$  car  $X_1$  est symétrique et  $-X_2 \stackrel{loi}{\cong} X_2$  car  $X_2$  est une gaussienne, donc symétrique également.

On a donc l'égalité :  $E \left[ \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq 5\}} \right] = E \left[ \mathbb{I}_{\{e^{-X_{1,i}} + e^{-X_{2,i}} \geq 5\}} \right]$ .

La méthode de Monte-Carlo par variables antithétiques donne ainsi l'estimateur :

$$p = \frac{1}{2n} \sum_{i=1}^n \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq 5\}} + \mathbb{I}_{\{e^{-X_{1,i}} + e^{-X_{2,i}} \geq 5\}}$$

#### Question 5.c :

On a :  $P[\sqrt{e^{X_1} + e^{X_2}} > k] = P[X_1 + X_2 > \log(k^2)]$ . La variable de contrôle de Monte-Carlo est donnée par :

$$\frac{1}{n} \sum_{i=1}^n g(X^i) - b(h(X^i) - E[h(X)])$$

Avec :  $h(X^i) = \mathbb{I}_{\{X_{1,i} + X_{2,i} > \log(k^2)\}}$  ;  $g(X^i) = \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq k\}}$  et  $b(X) = \frac{\text{cov}(h(X), g(X))}{\text{var}(h(X))}$

La méthode de Monte-Carlo par variable de contrôle donne ainsi l'estimateur :

$$p = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq k\}} - \frac{E \left[ \left( \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq k\}} - E \left[ \mathbb{I}_{\{e^{X_{1,i}} + e^{X_{2,i}} \geq k\}} \right] \right) \left( \mathbb{I}_{\{X_{1,i} + X_{2,i} > \log(k^2)\}} - E \left[ \mathbb{I}_{\{X_{1,i} + X_{2,i} > \log(k^2)\}} \right] \right) \right]}{\text{var} \left( \mathbb{I}_{\{X_{1,i} + X_{2,i} > \log(k^2)\}} - E \left[ \mathbb{I}_{\{X_{1,i} + X_{2,i} > \log(k^2)\}} \right] \right)}$$

On essaiera avec  $k = 5$  ou  $k = 2$

#### Question 5.d :

LOL

#### Question 6 :

Résultats :

## Second Cas

On s'intéresse désormais au calcul de cette intégrale :

$$\mathcal{I} = \int_{\mathbb{R}^2} \cos(y_1 y_2) \sin(y_1) e^{\sin(y_1 + y_2)} f_2(y_1, y_2) dy_1 dy_2$$

Avec  $(Y_1, Y_2) \sim f_2$ .

### Question 7.a :

La méthode de Monte-Carlo classique donne, avec  $(Y_1, Y_2) \sim f_2$  :

$$\begin{aligned} \mathcal{I} &= E[\cos(y_1 y_2) \sin(y_1) e^{\sin(y_1 + y_2)}] \\ \mathcal{I} &= \frac{1}{n} \sum_{i=1}^n \cos(y_{1,i} y_{2,i}) \sin(y_{1,i}) e^{\sin(y_{1,i} + y_{2,i})} \end{aligned}$$

### Question 7.b :

On a :

$$E[h(-Y_1, -Y_2)] = \int_{\mathbb{R}^2} h(-y_1, -y_2) \left\{ \cos^2(y_1) + \frac{1}{2} \sin^2(3y_2) \cos^4(y_1) \right\} e^{-\frac{1}{2} \left( \frac{y_1^2}{4} + y_2^2 \right)} dy_1 dy_2$$

Par changement de variable  $\begin{cases} s = -y_1 \\ t = -y_2 \end{cases}$ , on obtient la matrice hessienne  $J = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  de déterminant égal à 1.

On a donc :

$$E[h(S, T)] = \int_{\mathbb{R}^2} h(s, t) \left\{ \cos^2(s) + \frac{1}{2} \sin^2(3t) \cos^4(s) \right\} e^{-\frac{1}{2} \left( \frac{s^2}{4} + t^2 \right)} ds dt = E[h(Y_1, Y_2)]$$

Donc  $(-Y_1, -Y_2) \stackrel{\text{loi}}{\cong} (Y_1, Y_2)$ . L'estimateur de Monte Carlo par variables antithétiques s'écrit donc :

$$\mathcal{I} = \frac{1}{2n} \sum_{i=1}^n \cos(y_{1,i} y_{2,i}) \sin(y_{1,i}) e^{\sin(y_{1,i} + y_{2,i})} + \cos(-y_{1,i} - y_{2,i}) \sin(-y_{1,i}) e^{\sin(-y_{1,i} - y_{2,i})}$$

### Question 8 :

On observe que l'intervalle de confiance est meilleur avec les variables antithétiques. Cela n'est pas très surprenant vu que lorsque l'on utilise les variables antithétiques avec des sinus et des cosinus, toutes les variables sont réutilisables et donc nous avons deux fois plus de variables pour estimer qu'avec l'estimateur de Monte Carlo classique.

### III. Recyclage dans l'algorithme du rejet

#### Question 9 :

Pour tout  $i$ ,  $Z^{(i)}$  a pour densité  $\frac{Mg(y)-f(y)}{M-i}$ . **Justification ?**

#### Question 10 :

Soit  $U \sim g$  et  $V \sim g$  deux variables aléatoires indépendantes.

Montrons que les estimateurs  $\delta_1$  et  $\delta_2$  sont sans biais de  $E_f[h(X)]$ , i.e.  $E[\delta_1] = E[\delta_2] = E_f[h(X)]$ .

Comme  $U \sim g$ , on a  $E\left[h(U) \frac{f(U)}{g(U)}\right] = \int h(x) \frac{f(x)}{g(x)} g(x) dx = \int h(x) f(x) dx = E_f[h(X)]$ .

$$\begin{aligned} E[\delta_1] &= \frac{1}{t+1} \left( E\left[h(U) \frac{f(U)}{g(U)}\right] + \sum_{i=1}^t E[h(X^{(i)})] \right) \\ &= \frac{1}{t+1} (E_f[h(X)] + tE_f[h(X)]) \\ &= E_f[h(X)] \end{aligned}$$

Il reste à montrer que  $\delta_1 \xrightarrow{\mathbb{P}} E_f[h(X)]$ .

Montrons maintenant que  $E[\delta_2] = E_f[h(X)]$ . On a :

$$\sum_{i=1}^{N-t} h(Z^{(i)}) \frac{(M-1)f(Z^{(i)})}{Mg(Z^{(i)}) - f(Z^{(i)})} \xrightarrow{N \rightarrow +\infty} E\left[h(Z) \frac{(M-1)f(Z)}{Mg(Z) - f(Z)}\right]$$

Or

- $Z \sim \frac{Mg-f}{M-1}$  d'où  $E\left[h(Z) \frac{(M-1)f(Z)}{Mg(Z)-f(Z)}\right] = \int h(x) \frac{(M-1)f(x)}{Mg(x)-f(x)} \frac{Mg(x)-f(x)}{M-1} dx = \int h(x) f(x) dx = E_f[h(X)]$ .
- $V \sim g$  d'où  $E\left[h(V) \frac{f(V)}{g(V)}\right] = \int h(x) \frac{f(x)}{g(x)} g(x) dx = \int h(x) f(x) dx = E_f[h(X)]$ .

$$\begin{aligned} E[\delta_2] &= \frac{1}{N-t+1} \left( E\left[h(V) \frac{f(V)}{g(V)}\right] + \sum_{i=1}^{N-t} E\left[h(Z^{(i)}) \frac{(M-1)f(Z^{(i)})}{Mg(Z^{(i)}) - f(Z^{(i)})}\right] \right) \\ &= \frac{1}{N-t+1} (E_f[h(X)] + (N-t)E\left[h(Z) \frac{(M-1)f(Z)}{Mg(Z)-f(Z)}\right]) \\ &= \frac{1}{N-t+1} (E_f[h(X)] + (N-t)E_f[h(X)]) \\ &= E_f[h(X)] \end{aligned}$$

Ainsi on a bien  $E[\delta_1] = E[\delta_2] = E_f[h(X)]$ , donc les estimateurs  $\delta_1$  et  $\delta_2$  sont sans biais.

#### Question 11 :

On peut écrire  $\delta_1 = G(U, X^1, \dots, X^t)$  et  $\delta_2 = H(V, Z^1, \dots, Z^{N-t})$ . Or, on sait que :

- $U$  et  $V$  sont indépendantes

- $U$  et  $(X^1, \dots, X^t)$  sont indépendantes
- $U$  et  $(Z^1, \dots, Z^{N-t})$  sont indépendantes
- $V$  et  $(X^1, \dots, X^t)$  sont indépendantes
- $V$  et  $(Z^1, \dots, Z^{N-t})$  sont indépendantes
- $X_i$  et  $Z_i$  sont indépendantes car ce sont des tirages indépendants suivant le rejet.

D'où  $G(U, X^1, \dots, X^t)$  et  $H(V, Z^1, \dots, Z^{N-t})$  sont indépendantes.

Ainsi, les variables aléatoires  $\delta_1$  et  $\delta_2$  sont indépendantes.

Calculons la variance de  $\alpha\delta_1 + (1 - \alpha)\delta_2$ .

$$var(\alpha\delta_1 + (1 - \alpha)\delta_2) \stackrel{(*)}{=} \alpha^2 var(\delta_1) + (1 - \alpha)^2 var(\delta_2)$$

Calculons  $var(\delta_1)$  et  $var(\delta_2)$ .

$$\begin{aligned} var(\delta_1) &= \frac{1}{(t+1)^2} \left[ var\left(h(U) \frac{f(U)}{g(U)}\right) + \sum_{i=1}^t var\left(h(X^{(i)})\right) \right] \\ &= \frac{1}{(t+1)^2} \left[ var\left(h(U) \frac{f(U)}{g(U)}\right) + t \cdot var(h(X)) \right] \end{aligned}$$

Or

- $E\left[h^2(U) \frac{f^2(U)}{g^2(U)}\right] = \int h^2(y) \frac{f^2(y)}{g(y)} dy = \int h^2(y) \frac{f(y)}{g(y)} f(y) dy = E_f\left[\frac{h^2(X)}{g(X)} f(X)\right]$ , car  $X \sim f$ .
- $U \sim g$ , donc  $E\left[h(U) \frac{f(U)}{g(U)}\right] = E_f[h(X)]$

$$\begin{aligned} \text{Donc } var(\delta_1) &= \frac{1}{(t+1)^2} \left( E_f\left[\frac{h^2(X)}{g(X)} f(X)\right] - E_f[h(X)]^2 + t \cdot E_f[h^2(X)] - t \cdot E_f[h(X)]^2 \right) \\ &= \frac{1}{(t+1)^2} \left( E_f\left[\frac{h^2(X)}{g(X)} f(X)\right] + t \cdot E_f[h^2(X)] - (t-1) \cdot E_f[h(X)]^2 \right) \end{aligned}$$

$$\begin{aligned} var(\delta_2) &= \frac{1}{(N-t+1)^2} \left( var\left(h(V) \frac{f(V)}{g(V)}\right) + \sum_{i=1}^{N-t} var\left(h(Z^{(i)}) \frac{(M-1)f(Z^{(i)})}{Mg(Z^{(i)})-f(Z^{(i)})}\right) \right) \\ &= \frac{1}{(N-t+1)^2} \left( var\left(h(V) \frac{f(V)}{g(V)}\right) + (N-t) \cdot var\left(h(Z) \frac{(M-1)f(Z)}{Mg(Z)-f(Z)}\right) \right) \end{aligned}$$

Or

- $E\left[h^2(Z) \frac{(M-1)^2 f^2(Z)}{(Mg(Z)-f(Z))^2}\right] = \int h^2(y) \frac{(M-1)^2 f^2(y)}{Mg(y)-f(y)} dy = E_f\left[h^2(X) \frac{(M-1)f(X)}{Mg(X)-f(X)}\right]$
- $E\left[h(Z) \frac{(M-1)f(Z)}{Mg(Z)-f(Z)}\right] = E_f[h(X)]$

$$\begin{aligned} \text{Donc } var(\delta_2) &= \frac{1}{(N-t+1)^2} \left( E_f\left[\frac{h^2(X)}{g(X)} f(X)\right] - E_f[h(X)]^2 + (N-t) E_f\left[h^2(X) \frac{(M-1)f(X)}{Mg(X)-f(X)}\right] - (N-t) E_f[h(X)]^2 \right) \\ &= \frac{1}{(N-t+1)^2} \left( E_f\left[\frac{h^2(X)}{g(X)} f(X)\right] + (N-t) E_f\left[h^2(X) \frac{(M-1)f(X)}{Mg(X)-f(X)}\right] - (N-t+1) E_f[h(X)]^2 \right) \end{aligned}$$

Par la formule (\*), on trouve  $var(\alpha\delta_1 + (1 - \alpha)\delta_2)$ .