Projet Monte Carlo

Aminata Ndiaye Faniriana Rakoto Endor

December 2021

1 Exercice : Déambulation vers la sortie

1. (a) On pose F fonction de répartiton de la loi $\mathcal{N}(\mu, 1)$.

$$\begin{aligned}
\mathbb{P}(|X_1| > b-a) &= 1 - \mathbb{P}(|X_1| < b-a) \\
&= 1 - \mathbb{P}(a-b < X_1 < b-a) \\
&= 1 - [F(b-a) - F(a-b)] \\
&= 1 - F_{a,b}
\end{aligned}$$

Avec $F_{a,b} = F(b-a) - F(a-b) < 1$ par croissance de F. Notons que les ensembles $\{S = k\}_{k\in\mathbb{N}^*}$ sont disjoints. En effet, pour k, l des entiers tels que $k < l$ et en notant $c := b - a$, on a :

$$\begin{aligned}
\{S = k\} \cap \{S = l\} &= \left(\bigcap_{i=0}^{k} \{|X_i| < c\} \cap \{|X_k| > c\}\right) \cap \left(\bigcap_{i=0}^{l} \{|X_i| < c\} \cap \{|X_l| > c\}\right) \\
&= \left(\bigcap_{i=0,\ i\neq k}^{l} \{|X_i| < c\}\right) \cap \{|X_l| < c\} \cap \underbrace{(\{|X_k| < c\} \cap \{|X_k| > c\})}_{=\emptyset} \\
&= \emptyset
\end{aligned}$$

Ce qui justifie l'égalité suivante :

$$\begin{aligned}
\mathbb{P}(S \leq n) &= \sum_{k=1}^{n} \mathbb{P}(S = k) \\
&= \sum_{k=1}^{n} (1 - F_{a,b}) F_{a,b}^{k-1}
\end{aligned}$$

$$
\begin{aligned}
\mathbb{P}(S > n) &= 1 - \mathbb{P}(S \geq n) \\
&= 1 - \sum_{k=1}^{n} (1 - F_{a,b}) F_{a,b}^{k-1} \\
&= 1 - (1 - F_{a,b}) \sum_{k=1}^{n} F_{a,b}^{k-1} \\
&= 1 - (1 - F_{a,b}) \sum_{k=1}^{n-1} F_{a,b}^{k} \\
&= 1 - (1 - F_{a,b}) (\frac{1 - F_{a,b}^{n-1}}{1 - F_{a,b}}) \\
&= (F_{a,b})^{n-1} \underset{n \to +\infty}{\longrightarrow} 0
\end{aligned}
$$

(b) Notons que pour $n \in \mathbb{N}^*$ si $S \leq n$ alors $T \leq n$. En effet, si $S \leq n$ alors il existe un $m \leq n$ tel que $|X_m| > b - a$. Il y a maintenant deux cas possibles, soit $W_{m-1} \notin]a, b[$, dans ce cas là on a immédiatemment que $T \leq n$. Soit $W_{m-1} \in]a, b[$, dans ce cas comme $|X_m| > b - a$, on a que $W_m = W_{m-1} + X_m \notin]a, b[$, d'où $T \leq n$.
Ceci implique directement que $\{S \leq n\} \subset \{T \leq n\}$, donc $\mathbb{P}(S \leq n) \leq \mathbb{P}(T \leq n)$ d'où $\mathbb{P}(T \leq n) \geq 1 - \mathbb{P}(S > n) \xrightarrow[n \to \infty]{} 1$. Or $\{T \leq n\} \subset \{T < \infty\}$ donc $\mathbb{P}(T \leq n) \leq \mathbb{P}(T \leq \infty) \leq 1$ pour tout entier $n \in \mathbb{N}^*$ donc $\mathbb{P}(T < \infty) = 1$.

2. On peut exprimer δ comme :

$$
\begin{aligned}
\delta &= \mathbb{P}_\mu (W_T \geq b) \\
&= \mathbb{E}_\mu [\mathbb{1}_{W_T \geq b}]
\end{aligned}
$$

On a alors pour une suite $(X_{i,j})$ $i \in 1, \ldots n$ et $j \in \mathbb{N}^*$ i.i.d selon la loi $\mathcal{N}(\mu, 1)$, et une suite de temps d'atteintes $(T_i)_{i=1}^n$ définit comme $T_i = \inf\{m \geq 1 : \sum_{j=1}^m X_{i,j} \geq b\}$ l'estimateur de monte carlo classique :

$$
\hat{\delta}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\sum_{j=1}^{T_i} X_{i,j} \geq b}
$$

3. (a) Pour le temps de sortie T de l'intervalle $]a, b[$ définit précedemment que l'on fixe après simulation, on note f_T la densité du vecteur $(X_1, \ldots, X_T) \overset{iid}{\sim} \mathcal{N}(\mu, 1)$. Ainsi, par indépendance des X_i on a pour tout $x_1, \ldots, x_T \in \mathbb{R}$:

$$
\begin{aligned}
f_T(x_1, \ldots, x_T) &= \prod_{i=1}^{T} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \mu)^2\right) \\
&= \frac{1}{(2\pi)^{\frac{T}{2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{T}(x_i - \mu)^2\right)
\end{aligned}
$$

De la même manière, on définit $g_{T,\theta}$ la densité de $(X_1, \ldots, X_T) \overset{iid}{\sim} \mathcal{N}(\theta, 1)$ où T est le temps de sortie de $]a, b[$ de la marche aléatoire associé aux X_i. Ainsi, pour tout $x_1, \ldots, x_T \in \mathbb{R}$, on a :

$$
g_{T,\theta}(x_1, \ldots, x_T) = \frac{1}{(2\pi)^{\frac{T}{2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{T}(x_i - \theta)^2\right)
$$

On a ainsi :

$$\begin{aligned}
\delta &= \mathbb{E}_{g_{T,\theta}}\left[\mathbb{1}_{\sum_{i=1}^T X_i \geq b}\frac{f_T(X_1,\dots,X_T)}{g_{T,\theta}(X_1,\dots,X_T)}\right] \\
&= \mathbb{E}_{g_{T,\theta}}\left[\mathbb{1}_{\sum_{i=1}^T X_i \geq b}\frac{\exp\left(-\frac{1}{2}\sum_{i=1}^T (X_i-\mu)^2\right)}{\exp\left(-\frac{1}{2}\sum_{i=1}^T (X_i-\theta)^2\right)}\right]
\end{aligned}$$

Donc :

$$\delta = \mathbb{E}_{g_{T,\theta}}\left[\mathbb{1}_{\sum_{i=1}^T X_i \geq b}\exp\left(-\frac{1}{2}\sum_{i=1}^T\left(2X_i(\theta-\mu)+\mu^2-\theta^2\right)\right)\right] \tag{1}$$

Ce qui donne l'estimateur d'echantillonage préferentiel pour une suite $(X_{i,j})$ $i \in 1,\dots n$ et $j \in \mathbb{N}^*$ i.i.d selon la loi $\mathcal{N}(\theta,1)$, et une suite de temps d'atteintes $(T_i)_{i=1}^n$ associé à $(X_{i,j})_{j\in\mathbb{N}^*}$ pour $i \in \{1,\dots,n\}$ fixé :

$$\hat{\Delta}_n(\theta) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\sum_{j=1}^{T_i} X_{ij} \geq b}\exp\left(-\frac{1}{2}\sum_{j=1}^{T_i}\left(2X_{ij}(\theta-\mu)+\mu^2-\theta^2\right)\right)$$

(b) Pour $\theta = -\mu$, on trouve pour $(X_1,\dots,X_T) \overset{iid}{\sim} \mathcal{N}(-\mu,1)$:

$$\hat{\Delta}_n(-\mu) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\sum_{j=1}^{T_i} X_{ij} \geq b}\exp\left(2\mu\sum_{j=1}^{T_i} X_{i,j}\right)$$

D'après le cours on a alors :

$$Var\left(\hat{\Delta}_n(-\mu)\right) = \frac{1}{n}\left(\mathbb{E}_{g_{T,-\mu}}\left[\mathbb{1}^2_{W_T\geq b}\exp\left(4\mu\sum_{j=1}^T X_j\right)\right] - \delta^2\right)$$

et :

$$Var\left(\hat{\delta}_n\right) = \frac{1}{n}\left(\mathbb{E}_{f_T}\left[\mathbb{1}^2_{W_T\geq b}\right] - \delta^2\right)$$

Donc on a que $Var(\hat{\Delta}_n(-\mu)) \leq Var(\hat{\delta}_n)$ si et seulement si :

$$\mathbb{E}_{g_{T,-\mu}}\left[\mathbb{1}_{W_T\geq b}\exp\left(4\mu\sum_{j=1}^T X_j\right)\right] \leq \mathbb{E}_{f_T}\left[\mathbb{1}_{W_T\geq b}\right] \tag{2}$$

Calculons le membre de gauche de l'inégalité et exprimons le comme une espérance par rapport à la mesure de densité f_T :

$$
\begin{aligned}
\mathbb{E}_{g_{T,-\mu}}\left[\mathbb{1}_{W_T\geq b}\exp\left(4\mu W_T\right)\right] &= \int_{\mathbb{R}^T} \mathbb{1}_{\sum_{j=1}^T x_j} \exp\left(4\mu\sum_{j=1}^T x_j\right) g_{T,-\mu}(x_1,\ldots,x_T)dx \\
&= \int_{\mathbb{R}^T} \mathbb{1}_{\sum_{j=1}^T x_j} \exp\left(4\mu\sum_{j=1}^T x_j\right) \frac{1}{2\pi^{\frac{T}{n}}}\exp\left(-\frac{1}{2}\sum_{j=1}^T (x_j+\mu)^2\right)dx \\
&= \int_{\mathbb{R}^T} \mathbb{1}_{\sum_{j=1}^T x_j} \exp\left(4\mu\sum_{j=1}^T x_j\right) \frac{1}{2\pi^{\frac{T}{n}}}\exp\left(-\frac{1}{2}\sum_{j=1}^T (x_j-\mu+2\mu)^2\right)dx \\
&= \int_{\mathbb{R}^T} \mathbb{1}_{\sum_{j=1}^T x_j} \exp\left(2\mu\sum_{j=1}^T x_j\right) \frac{1}{(2\pi)^{\frac{T}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^T (x_i-\mu)^2\right)dx \\
&= \int_{\mathbb{R}^T} \mathbb{1}_{\sum_{j=1}^T x_j} \exp\left(2\mu\sum_{j=1}^T x_j\right) f_T(x_1,\ldots,x_T)dx \\
&= \mathbb{E}_{f_T}\left[\mathbb{1}_{W_T\geq b}\exp\left(2\mu W_T\right)\right]
\end{aligned}
$$

Or dans cette expression $W_T \geq b > 0$ car on se restreint à integrer sur $\{W_T \geq b\}$, donc comme $\mu < 0$ on a que $\exp(2\mu W_T) < 1$ et par croissance de l'espérance par rapport à la mesure de densité f_T dans la relation (2), on a bien que $Var\left(\hat{\Delta}_n(-\mu)\right) \leq Var\left(\hat{\delta_n}\right)$.

4. D'après l'égalité (1) on a :

$$\delta = \mathbb{E}_{g_{T,-\mu}}\left[\mathbb{1}_{W_T\geq b}\exp\left(2\mu W_T\right)\right]$$

Ainsi, comme $\mu < 0$ et qu'on intègre sur $\{W_T \geq b\}$ on a par croissance de l'exponentielle que $2\mu W_T \leq 2\mu b$. Ce qui donne :

$$
\begin{aligned}
\delta &\leq \mathbb{E}_{g_{T,-\mu}}\left[\mathbb{1}_{W_T\geq b}\exp\left(2\mu b\right)\right] \\
&\leq \mathbb{E}_{g_{T,-\mu}}\left[\exp\left(2\mu b\right)\right] \\
&= \exp(2\mu b)
\end{aligned}
$$

En prenant $\mu = -1$ et $b = 5$, on trouve $\exp(2\mu b) \simeq 5\cdot 10^{-5}$ donc la majoration $\delta \leq 5\cdot 10^{-5}$.
Trouvons maintenant l'entier $n \in \mathbb{N}^*$ tel l'écart entre l'estimateur et l'estimé soit faible. Comme δ est de l'ordre de 10^{-5} on peut prendre une précision ε de cet ordre pour que l'écart soit minime. En effet, comme le developpement décimal de δ commence au moins à partir de 10^{-5}, il est souhaitable de prendre $\varepsilon = 10^{-5}$. Prendre ε un ordre de grandeur en dessous à 10^{-6} multipliera considérablement le nombre de variables à simuler et prendre ε un ordre de grandeur au dessus à 10^{-4} risque de donner une approximation grossière de δ.
On va maintenant utiliser le théorème centrale limite pour trouver le bon n. Considérons une suite de variables aléatoires $(\mathbb{1}_{W_{T_i}\geq b})_{i\geq 1} \overset{iid}{\sim} \mathcal{B}(\delta)$. Cette suite est de carrée integrable car elle admet une variance de $\delta(1-\delta)$, on peut donc appliquer le théorème centrale limite :

$$\sqrt{n}(\hat{\delta}_n - \delta) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \delta(1-\delta))$$

Comme $\delta(1-\delta) > 0$ on obtient avec $0 < \alpha < 1$ et $q_{1-\frac{\alpha}{2}}$ le quantile de la loi $\mathcal{N}(0,1)$ d'ordre $1-\frac{\alpha}{2}$ ainsi que le théorème de Slutsky :

$$\lim_{n\to\infty}\mathbb{P}\left(\sqrt{\frac{n}{\delta(1-\delta)}}\left|\hat{\delta}_n - \delta\right| \leq q_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

Ce qui est équivalent à :

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\hat{\delta}_n - \delta\right| > q_{1-\frac{\alpha}{2}}\sqrt{\frac{\delta(1-\delta)}{n}}\right) = \alpha$$

En posant $\varepsilon = q_{1-\frac{\alpha}{2}}\sqrt{\frac{\delta(1-\delta)}{n}}$ on obtient $n = \delta(1-\delta)\frac{q_{1-\frac{\alpha}{2}}{}^2}{\varepsilon^2}$. En faisant l'application numérique avec $\delta = 5\cdot 10^{-5}$, $\alpha = 0.05$ et $\varepsilon = 10^{-5}$ on a finalement $n \simeq 1.9\cdot 10^6$ qui est de l'ordre du n attendu.

5. On trouve avec le seed 1, $\hat{\delta}_n = 1.5\cdot 10^{-5}$ et $\hat{\Delta}_n(1) = 1.453379\cdot 10^{-5}$ avec leurs variances respéctifs $Var(\hat{\delta}_n) = 1.499979\cdot 10^{-5}$ et $Var(\hat{\Delta}_n(1)) = 1.617645\cdot 10^{-10}$ et coûts moyens respectifs pour 100 itérations avec microbenchmark $cout\_mcc = 480580201$ et $cout\_ep = 763326052$. Ce qui donne une efficacité relative :
$$\frac{Var(\hat{\delta}_n)\ cout\_mcc}{Var(\hat{\Delta}_n(1))\ cout\_ep} = 58379.13$$
On conclut donc que la simulation par echantillonage préférentiel est plus efficace en générale. Notons tout de même qu'en terme de coût, la simulation par la méthode de monte carlo classique est plus efficace même si les deux coûts sont du même ordre de grandeur. Mais c'est en terme de variance que la methode de simulation par echantillonage préférentiel est meilleure avec une variance de 5 ordre de grandeur plus petite.

2 Exercice : Recettes de financier

Partie I

1. (a) On a $W_T = \sum_{k=1}^T X_k$ avec $(X_n)_{n\in\mathbb{N}} \overset{iid}{\sim} \mathcal{N}(\mu, 1)$
W_T est donc une combinaison linéaire de variables aléatoires gaussiennes indépendantes. W_T suit donc une loi normale.

$$\mathbb{E}[W_T] = \mathbb{E}[\sum_{k=1}^T X_k] = \sum_{k=1}^T \mathbb{E}[X_k] = T\mathbb{E}[X_1] = T\mu$$

$$Var(W_T) = Var(\sum_{k=1}^T X_k) = \sum_{k=1}^T Var(X_k) = TVar(X_1) = T$$

D'où $W_T \sim \mathcal{N}(T\mu, T)$
On en déduit l'estimateur de Monte Carlo classique $\hat{\delta}_n$:

$$\hat{\delta}_n = \frac{1}{n}\sum_{i=1}^n max\{\lambda exp(-\sigma Y_i) - K, 0\}\ (Y_i)_{1\le i\le n} \sim \mathcal{N}(T\mu, T)$$

(b) L'erreur quadratique moyenne relativement à δ :

$$\mathbb{E}[(\hat{\delta}_n - \delta)^2] = Var(\hat{\delta}_n) + Biais(\hat{\delta}_n, \delta)^2$$

Or $\hat{\delta}_n$ est un estimateur de Monte Carlo de δ, il est donc sans biais. Donc :

$$\mathbb{E}[(\hat{\delta}_n - \delta)^2] = Var(\hat{\delta}_n)$$

Application numérique (voir code)

$$Var(\hat{\delta}_n) = 33,73$$
$$\hat{\delta}_n = 5.35$$

2. (a) On peut considérer l'application A suivante qui laisse la loi ν invariante :

$$A : W \longrightarrow 2T\mu - W$$

On en déduit l'estimateur de δ par la méthode de la variable antithétique :

$$\hat{\delta}_n(A) = \frac{1}{n}\sum_{k=1}^{n} \frac{max\{\lambda exp(-\sigma Y_i) - K, 0\} + max\{\lambda exp(-\sigma A(Y_i)) - K, 0\}}{2}$$

De plus comme A est une transformation décroissante et

$$h: \ X \longmapsto max\{\lambda exp(-\sigma X) - K, 0\}$$

est décroissante, on a donc d'après le corollaire du cours que :

$$Cov(h(Y_i), h(A(Y_i))) < 0$$

d'où :

$$Cor(h(Y_i), h(A(Y_i))) < 0$$

On a :

$$\begin{aligned} Var(\hat{\delta}_n) &= Var\left(\frac{1}{n}\sum_{k=1}^{n} h(Y_i)\right) \\ &= \frac{Var(h(X_1)}{n} \end{aligned}$$

$$\begin{aligned} Var(\hat{\delta}_n(A)) &= Var\left(\frac{1}{n}\sum_{k=1}^{n} \frac{h(Y_i) + h(A(Y_i))}{2}\right) \\ &= \frac{1}{4n}[Var(h(X_1) + h(A(X_1)))] \\ &= \frac{1}{4n}[Var(h(X_1)) + Var(h(A(X_1))) + 2Cov(h(X_1), h(A(X_1))] \\ &= \frac{Var(h(X_1)) + Cov(h(X_1), h(A(X_1))}{2n} \\ &= \frac{Var(\hat{\delta}_n)[1 + Cor(h(Y_i), h(A(Y_i)))]}{2} \end{aligned}$$

On en conclut donc que :

$$Var(\hat{\delta}_n(A)) \leq Var(\hat{\delta}_n)$$

(b)

$$\mathbb{E}[(\hat{\delta}_n(A) - \delta)^2] = Var(\hat{\delta}_n(A)) = 4.45$$

3. La méthode de la variable antithétique est 6 fois plus efficace que la méthode de Monte carlo classique.

Partie II

4. (a) Estimateur de Monte Carlo classique de $\mathcal{I}$:

$$\hat{\mathcal{I}}_n = \frac{1}{n}\sum_{k=1}^{T} max(\lambda_1 exp(-\sigma_1 X_{1,i}) + \lambda_2 exp(-\sigma_2 X_{2,i}) - K, 0)$$

(b)

$$\hat{\mathcal{I}}_n = 8.72$$

Intervalle de confiance asymptotique bilatère au niveau 90% de l'estimateur de Monte Carlo classique :

$$[8.514303; 8.934779]$$

5. (a) On utilise l'expression de la fonction génératrice des moments de la loi normale $\mathcal{N}(0,1)$:

$$\mathbb{E}[e^{tX}] = exp\left(\frac{t^2}{2}\right)$$

Avec l'application numérique on obtient (voir code) :

$$\mathbb{E}[e^{-\sigma_1 X}] = 1.13$$

$$\mathbb{E}[e^{-\sigma_2 X}] = 1.05$$

$$\mathbb{E}[\lambda_1 e^{-\sigma_1 X_1} + \lambda_2 e^{-\sigma_2 X_2}] = 73.81$$

(b) Nous avons choisi la variable de contrôle h_0 qui maximise en valeur absolue la corrélation entre $h(X_1, X_2)$ et $h_0(X_1, X_2)$ avec :

$$h : (X_1, X_2) \longrightarrow max(\lambda_1 exp(-\sigma_1 X_1) + \lambda_2 exp(-\sigma_2 X_2) - K, 0)$$

On choisit (voir code) :

$$h_0 : (X_1, X_2) \longrightarrow \lambda_1 e^{-\sigma_1 X_1} + \lambda_2 e^{-\sigma_2 X_2}$$

On choisit le b optimal vu en cours :

$$b = \frac{Cov(h(X_1), h_0(X_1))}{Var(h_0(X_1))}$$

On utilise une période de chauffe de 2000 observations (voir code) pour estimer convenablement ce b optimal. Finalement on obtient :

$$\hat{\mathcal{I}}_n(b) = 8.53$$

Intervalle de confiance asymptotique bilatère au niveau 90% de l'estimateur avec la méthode de variables de contrôle :

$$[8.46256; 8.589278]$$

6. On obtient une efficacité relative de 10.43 .L'estimateur avec la méthode de variables de contrôle est 10 fois plus efficace que l'estimateur de Monte carlo classique. De plus on obtient un intervalle de confiance asymptotique plus petit.

Partie III

7. Estimation par la méthode de Monte carlo classique :

$$\hat{p}_n = 0,0595$$

$$Var(\hat{p}_n) = 0,012$$

8. $\mathbf{Z}$ et $\mathbf{X}$ sont deux vecteurs gaussiens standards indépendants $< \mathbf{X}, u > u + \mathbf{Z} - < \mathbf{Z}, u > u$ est une combinaison linéaire de deux vecteurs gaussiens indépendants et suit donc une loi normale dont on détermine l'espérance et la variance par les calculs suivants :

$$\begin{aligned}
\mathbb{E}[< \mathbf{X}, u > u + \mathbf{Z} - < \mathbf{Z}, u > u] &= \mathbb{E}\left[u \sum_{i=1}^{d} X_i u_i + \mathbf{Z} - u \sum_{i=1}^{d} Z_i u_i\right] \\
&= \mathbb{E}\left[u \sum_{i=1}^{d} (X_i - Z_i) u_i + \mathbf{Z}\right] \\
&= u \sum_{i=1}^{d} u_i \Bigg(\underbrace{\mathbb{E}[X_i]}_{=0} - \underbrace{\mathbb{E}[Zi]\Bigg)}_{=0} + \underbrace{\mathbb{E}[\mathbf{Z}]}_{=0_d} \\
&= 0_d
\end{aligned}$$

$$\begin{aligned}
Var\Bigg(< \mathbf{X}, u > u + \mathbf{Z} - < \mathbf{Z}, u > u\Bigg) &= Var\Bigg(\mathbf{X}^T.uu + \mathbf{Z}^T.uu\Bigg) \\
&= Var\Bigg(\mathbf{X}^T.uu\Bigg) + Var\Bigg(\mathbf{Z}^T.uu\Bigg) \\
&= uVar(\sum_{i=1}^{d} X_i u_i) u^T + var\Bigg(\mathbf{Z} - \mathbf{Z}^T.uu\Bigg) \\
&= u\left[\sum_{i=1}^{d} u_i^2 Var(X_i)\right] u^T + Var(\mathbf{Z} - \mathbf{Z}^T.uu) \\
&= u.u^T \underbrace{\|u\|^2}_{=1} + Var(\mathbf{Z}) + Var(\mathbf{Z}^T.uu) + 2Cov(\mathbf{Z}, -\mathbf{Z}^T.uu) \\
&= u.u^T + \mathbb{I}_d + u.u^T \|u\|^2 - 2Cov(\mathbf{Z}, \sum_{i=1}^{d} u_i Z_i u) \\
&= u.u^T + \mathbb{I}_d + u.u^T - 2\sum_{i=1}^{d} u_i Cov(\mathbf{Z}, Z_i u) \\
&= u.u^T + \mathbb{I}_d + u.u^T - 2u.u^T \\
&= \mathbb{I}_d
\end{aligned}$$

9. D'après la question précédente $\mathbf{X}$ suit la même loi que $< \mathbf{X}, u > u + \mathbf{Z} - < \mathbf{Z}, u > u$ et $(D_1, ..., D_L)$ est une partition de $\mathbb{R}$, on a donc :

$$
\begin{aligned}
p &= \mathbb{E}\left[max\left\{\frac{1}{d}\sum_{i=1}^{d} exp(\sigma(< \mathbf{X}, u > u_i + \mathbf{Z}_i - < \mathbf{Z}, u > u_i) - K, 0\right\}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[max\left\{\frac{1}{d}\sum_{i=1}^{d} exp(\sigma(< \mathbf{X}, u > u_i + \mathbf{Z}_i - < \mathbf{Z}, u > u_i) - K, 0\right\}\middle| < \mathbf{X}, u >\right]\right] \\
&= \sum_{l=1}^{L} \mathbb{P}[< \mathbf{X}, u >\in D_l]\mathbb{E}\left[max\left\{\frac{1}{d}\sum_{i=1}^{d} exp(\sigma(< \mathbf{X}, u > u_i + \mathbf{Z}_i - < \mathbf{Z}, u > u_i) - K, 0\right\} \mid < \mathbf{X}, u >\in D_l\right] \\
&= \sum_{l=1}^{L} \mathbb{P}[< \mathbf{X}, u >\in D_l]\mathbb{E}\left[max\left\{\frac{1}{d}\sum_{i=1}^{d} exp(\sigma(Y_l u_i + \mathbf{Z}_i - < \mathbf{Z}, u > u_i) - K, 0\right\}\right]
\end{aligned}
$$

10. (a) $< \mathbf{X}, u >$ est une combinaison linéaire d'un vecteur gaussien et suit donc une loi normale. De plus :

$$
\begin{aligned}
\mathbb{E}[< \mathbf{X}, u >] &= \mathbb{E}\left[\sum_{i=1}^{d} X_i u_i\right] \\
&= \sum_{i=1}^{d} \mathbb{E}[X_i] u_i \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
Var(< \mathbf{X}, u >) &= Var(\sum_{i=1}^{d} X_i u_i) \\
&= \sum_{i=1}^{d} Var(X_i) u_i^2 \\
&= \|u\|^2 \\
&= 1
\end{aligned}
$$

$< \mathbf{X}, u >$ suit donc une loi $\mathcal{N}(0,1)$. On utilise la fonction quantile de la loi $\mathcal{N}(0,1)$ pour obtenir une partition $(D_1, ..., D_L)$ telle que pour $l \in [\![1, L]\!], \mathbb{P}[< \mathbf{X}, u >\in D_l] = \frac{1}{L}$.
On pose donc $\forall l \in [\![1, L]\!]$, :

$$D_l = \left[\overleftarrow{F}\left(\frac{l-1}{L}\right); \overleftarrow{F}\left(\frac{l}{L}\right)\right[$$

Avec $\overleftarrow{F}$ la fonction quantile de la $\mathcal{N}(0,1)$.

(b) Nous justifions graphiquement l'utilisation de l'approximation polynomiale de l'inverse généralisé de la loi $N(0,1)$

$$\hat{p}_n(u) = 0,057$$

$$Var(\hat{p}_n(u)) = 0,00023$$

11. L'estimateur avec la méthode de stratification avec allocation proportionnelle est plus efficace que l'estimateur de Monte carlo classique. Il y a une erreur dans notre code pour l'estimateur avec la méthode de stratification que nous n'avons pas su détecter qui sous évalue la variance de la méthode et donc surévalue l'efficacité relative de la méthode.