Pricing des options asiatiques.

Tony Lelièvre

Juin 2000

Le but de ce rapport est de donner quelques méthodes pour le pricing des options asiatiques, avec des programmes qui ont été testés. J'ai suivi deux pistes : l'une donnée par [LT] et qui est basée sur des méthodes de Monte-Carlo. J'ai prolongé les simulations en adaptant une méthode de réduction de variance efficace dans le cas où la valeur initiale du sous-jacent est plus petite (dans le cas d'un call) que le strike. L'autre piste est celle de la résolution d'une EDP qui est obtenue après des changements de variables astucieux (cf. [RS]) et qui permet également (en théorie du moins...) de trouver le prix d'une option asiatique. J'ai donc cherché à programmer des méthodes de discrétisations de cette EDP, sans trop de succès.

De manière générale, le pricing des options asiatiques semble beaucoup plus simple avec des méthodes numériques probabilistes.

La première partie de ce rapport donne le contexte général et précise les notations. On donne ensuite des méthodes de Monte Carlo pour pricer l'option, puis des discrétisations de l'EDP donnée dans [RS]. En annexe, on donne le listing des programmes : en C pour les méthodes de Monte Carlo et sous MatLab pour la discrétisation de l'EDP.

1 Options asiatiques

On travaille dans le modèle de Black et Scholes. On a donc un sous-jacent risqué S_t et un sous-jacent sans risque S_t^0 et ils vérifient :

$$\begin{aligned} dS_t &= S_t(\mu\, dt + \sigma\, dB_t) \\ dS_t^0 &= rS_t^0\, dt \end{aligned}$$

où μ (le rendement), σ (la volatilité $\sigma > 0$) et r (taux d'emprunt, taux d'intérêt $r > 0$) sont trois constantes et B_t un brownien standard. Pour pricer l'option, on travaille en fait sous la probabilité risque neutre, pour laquelle $W_t = B_t + \frac{\mu - r}{\sigma} t$ est un brownien (cf. [LL]). Sous cette probabilité, le prix actualisé $\tilde{S}_t = \exp(-rt)S_t$ est une martingale. On a alors,

$$dS_t = S_t(r\, dt + \sigma\, dW_t)$$

qui admet pour solution

$$S_t = S_0 \exp\left((r - \frac{\sigma^2}{2})t + \sigma W_t \right).$$

Une option asiatique est une option dont le payoff dépend de la moyenne du sous-jacent $I = \frac{1}{T}\int_0^T S_u\, du$. Dans la suite, on s'intéresse aux calls et aux puts à strike fixe dont la fonction de payoff est respectivement $\phi(I) = (I - K)^+$ et $\phi(I) = (K - I)^+$. On cherche donc à calculer le prix de cette option, qui est donné par $P = e^{-rT}\mathbb{E}(\phi(I))$.

2 Méthodes de Monte Carlo

Pour trouver le prix d'une option asiatique, il faut donc simuler la variable aléatoire $I = \frac{1}{T}\int_0^T S_0 \exp\left((r - \frac{\sigma^2}{2})t + \sigma W_t\right) dt$, ou plus exactement $\phi(I)$. Le problème est qu'on ne connaît pas explicitement la loi de I. On sait simplement simuler la trajectoire d'un brownien W_t, en des temps discrets. Il faut donc discrétiser en temps l'intégrale dans I puis simuler la trajectoire.

Dans la suite, on note N le nombre de discrétisation en temps (le pas en etmps est $\delta t = T/N$) et M le nombre de réalisation de la variable aléatoire que l'on effectue pour approcher la moyenne (Monte Carlo). On aura donc deux types d'erreur dans l'approximation : une liée à la discrétisation en temps (qui sera typiquement en $O\left(\frac{1}{N}\right)$) et une liée à la méthode de Monte Carlo (qui sera typiquement en $O\left(\frac{\theta}{\sqrt{M}}\right)$, θ^2 étant la variance). La complexité de tous ces algorithmes est en $O(NM)$.

Sauf mention explicite, les simulations numériques sont faites avec les valeurs suivantes : $S_0 = 100$, $K = 100$, $\sigma = 0.3$, $r = 0.02$, $T = 1$ an.
La valeur exacte dans ce cas est environ 7.3078 (intervalle de confiance $[7.3063, 7.3094]$) : ce résultat a été obtenu avec un des schémas les plus performants (SchemaRed3) et avec $N = 1000$ (pas de temps $\delta t = 0.001$) et $M = 10^6$ (nombre de trajectoires engendrées).

2.1 Schémas de base

Schéma 1

Une manière simple d'approcher I est de l'écrire sous la forme $I_1 = \frac{1}{N}\sum_{k=0}^{N-1} S_{t_k}$. Pour calculer les S_{t_k}, il suffit de calculer les W_{t_k}, ce qui est facilement réalisable car les accroissements $W_{t_k} - W_{t_{k-1}}$, $k = 1..N-1$ sont des variables aléatoires gaussiennes centrées de variance δt i.i.d.. Dans le programme donné en annexe, il s'agit de la procédure Schema1. On écrit donc que $e^{-rT}\mathbb{E}(\phi(I)) \simeq \frac{1}{M}\sum_{i=1}^M X_i$ avec $X_i = e^{-rT}\phi(I_1^i)$ où I_1^i est une réalisation de la variable aléatoire I_1. La variance est également calculée numériquement en utilisant la formule $Var(X) \simeq \frac{1}{M-1}\left(\sum_{i=1}^M X_i^2 - M\left(\frac{1}{M}\sum_{i=1}^M X_i\right)^2\right)$. On peut donc donner un intervalle de confiance en utilisant cette valeur numérique. Les résultats sont donnés dans les tableaux 1 et 2.

Remarque : Les calculs ont été faits sur station. On utilise un générateur de nombre aléatoire, auquel on fournit une valeur initiale à chaque lancement du programme.

Schéma 2

Une méthode pour essayer de mieux approcher I est de remarquer qu'une fois la trajectoire W_{t_k} simulée (de manière exacte), on connaît la loi de $\mathcal{L}(W_u | W_{t_k} = x, W_{t_{k+1}} = y)$ pour tout $u \in (t_k, t_{k+1})$. Ce résultat repose sur le lemme suivant.

Lemme 1 *Soit (x, y) un vecteur gaussien avec x à valeurs dans $\mathbb{R}$ et y à valeurs dans $\mathbb{R}$. On pose $\theta_1 = \mathbb{E}(x^2) - \mathbb{E}(x)^2$, $\theta_2 = \mathbb{E}(y^2) - \mathbb{E}(y)^2$ et $\theta_3 = \mathbb{E}(xy) - \mathbb{E}(x)\mathbb{E}(y)$. La matrice de variance covariance de (x, y) est $\Gamma = \begin{bmatrix} \theta_1 & \theta_3 \\ \theta_3 & \theta_2 \end{bmatrix}$. On pose $\lambda = \frac{\theta_3}{\theta_2}$. Alors,*

$$\mathbb{E}(x|y) = \mathbb{E}(x) + \lambda(y - \mathbb{E}(y)).$$

De plus, si on pose $z = (x - \mathbb{E}(x)) - \lambda(y - \mathbb{E}(y))$, z est une gaussienne centrée et de variance $\theta = \theta_1 - \frac{\theta_3^2}{\theta_2}$. On a donc $x = \mathbb{E}(x) + \lambda(y - \mathbb{E}(y)) + z$ avec z indépendant de y. La loi *de x conditionnellement en y est donc celle d'une gaussienne centrée en $\mathbb{E}(x)+\lambda(y-\mathbb{E}(y))$ et de variance θ.*

Preuve : (z, y) est clairement un vecteur gaussien. Pour vérifier que z est indépendant de y, il suffit donc de calculer la covariance $\mathbb{E}(zy)$.

$$\begin{aligned}\mathbb{E}(zy) &= \mathbb{E}(z(y-\mathbb{E}(y))) \\ &= \theta_3 - \lambda\theta_2 \\ &= 0\end{aligned}$$

On a donc :

$$\begin{aligned}\mathbb{E}(x|y) &= \mathbb{E}((z+\mathbb{E}(x)+\lambda(y-\mathbb{E}(y)))|y) \\ &= \mathbb{E}(z)+\mathbb{E}(x)+\lambda(y-\mathbb{E}(y)) \\ &= \mathbb{E}(x)+\lambda(y-\mathbb{E}(y))\end{aligned}$$

Enfin, calculons la variance de z.

$$\begin{aligned}\mathbb{E}(z^2) &= \mathbb{E}\left(((x-\mathbb{E}(x))-\lambda(y-\mathbb{E}(y)))((x-\mathbb{E}(x))-\lambda(y-\mathbb{E}(y)))\right) \\ &= \theta_1 - \lambda\theta_3 - \theta_3\lambda + \lambda\theta_2\lambda \\ &= \theta_1 - \frac{\theta_3^2}{\theta_2}\end{aligned}$$

$\diamondsuit$

Pour trouver la loi de W_u , on applique le lemme précédent avec $x = W_u$ et $y = W_{t_{k+1}}$, en faisant démarrer le brownien en $W_{t_k} = x_k$ i.e. $\mathbb{E}(x) = x_k$ et $\mathbb{E}(y) = x_k$. On a donc $\theta_1 = (u - t_k)$, $\theta_2 = \delta t$ et $\theta_3 = \theta_1$ (cf. l'indépendance des accroissements des trajectoires) d'où $\lambda = \frac{u-t_k}{\delta t}$ et $\theta = \frac{(u-t_k)(t_{k+1}-u)}{\delta t}$. On obtient alors, pour $u \in (t_k, t_{k+1})$, $\mathcal{L}(W_u|W_{t_k} = x_k, W_{t_{k+1}} = x_{k+1}) = \mathcal{N}\left(\frac{t_{k+1}-u}{\delta t}x_k + \frac{u-t_k}{\delta t}x_{k+1}, \frac{(u-t_k)(t_{k+1}-u)}{\delta t}\right)$.

On utilise également le lemme suivant (obtenu par un simple calcul).

Lemme 2 *Si G est une gaussienne centrée réduite, alors*

$$\mathbb{E}(e^{\mu+\sigma G}) = e^{\mu+\sigma^2/2}$$

On approche alors I par sa projection sur l'ensemble des variables aléatoire bornée et mesurable par rapport à la tribu engendrée par $W_{t_0}...W_{t_{N-1}}$:

$$\begin{aligned}\mathbb{E}(I|W_{t_0}...W_{t_{N-1}}) &= \frac{S_0}{T}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left(\exp\left((r-\frac{\sigma^2}{2})u+\sigma W_u\right)|W_{t_k}, W_{t_{k+1}}\right) du \\ &= \frac{S_0}{T}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \exp\left((r-\frac{\sigma^2}{2})u+\sigma\left(\frac{t_{k+1}-u}{\delta t}W_{t_k}+\frac{u-t_k}{\delta t}W_{t_{k+1}}\right)+\right. \\ &\quad \left.\frac{\sigma^2}{2}\frac{(u-t_k)(t_{k+1}-u)}{\delta t}\right) du \\ &= \frac{1}{T}\sum_{k=0}^{N-1} S_{t_k}\int_{t_k}^{t_{k+1}} \exp\left(r(u-t_k)+\sigma\frac{u-t_k}{\delta t}(W_{t_{k+1}}-W_{t_k})-\right. \\ &\quad \left.\frac{\sigma^2}{2}\frac{(u-t_k)^2}{\delta t}\right) du \\ &\simeq \frac{1}{T}\sum_{k=0}^{N-1} S_{t_k}\left(\delta_t + r\frac{(\delta t)^2}{2}+\sigma\frac{\delta t}{2}(W_{t_{k+1}}-W_{t_k})\right)\end{aligned}$$

la dernière estimation correspondant à un développement de Taylor de l'exponentielle sous l'intégrale. On approche donc I par

$$I_2 = \frac{1}{N} \sum_{k=0}^{N-1} S_{t_k} \left(r\frac{\delta t}{2} + \frac{\sigma}{2}(W_{t_{k+1}} - W_{t_k}) \right).$$

Dans le programme donné en annexe, il s'agit de la procédure schema2. Les résultats sont donnés dans les tableaux 1 et 2.

Remarque : Ce schéma revient à approcher l'intégrale par la méthode des trapèzes. On peut en effet montrer que (cf. [LT]) $\mathbb{E}\left(\left(I_2 - \frac{1}{T}\sum_{k=0}^{N-1} \delta t \frac{S_{t_k}+S_{t_{k+1}}}{2}\right)^2\right) = O\left(\frac{1}{N^3}\right)$.

Schéma 3

Le dernier schéma repose sur le fait que l'on connaît explicitement la loi de $\int_0^T W_u\,du$.

Lemme 3 *On sait que pour $u \in (t_k, t_{k+1})$, $\mathcal{L}(W_u | W_{t_k} = x_k, W_{t_{k+1}} = x_{k+1}) = \mathcal{N}(M(u), \Theta(u))$ avec $M(u) = \frac{t_{k+1}-u}{\delta t}x_k + \frac{u-t_k}{\delta t}x_{k+1}$ et $\Theta(u) = \frac{(u-t_k)(t_{k+1}-u)}{\delta t}$. On peut alors explicitement trouver la loi de $\int_{t_k}^{t_{k+1}} W_u\,du$ sachant que $W_{t_k} = x_k$ et $W_{t_{k+1}} = x_{k+1}$. Il s'agit d'une loi normale de moyenne $m(x_k, x_{k+1}) = \frac{x+y}{2}\delta t$ et de variance $\frac{h^3}{6}$.*

Preuve : La moyenne est donnée par $\int_{t_k}^{t_{k+1}} M(u)\,du$ ce qui donne $m(x_k, x_{k+1})$. Pour calculer la variance, on utilise le fait que le brownien est à acroissements indépendants :

$$\begin{aligned}
\mathbb{E}\left(\left(\int_{t_k}^{t_{k+1}} W_u - M(u)\,du\right)^2\right) &= \mathbb{E}\left(\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}} (W_u - M(u))(W_v - M(v))\,dv\,du\right) \\
&= \mathbb{E}\left(\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}} (W_{(u\wedge v)} - M(u\wedge v))\,dv\,du\right) \\
&= \int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}} \Theta(u\wedge v)\,dv\,du \\
&= 2\int_{t_k}^{t_{k+1}}\int_{t_k}^{u} \Theta(v)\,dv\,du \\
&= \frac{h^3}{6}
\end{aligned}$$

où $u \wedge v = \inf(u, v)$. ◊

Remarque : On montre de même que $\mathcal{L}(\frac{1}{T}\int_0^T W_u\,du) = \mathcal{N}(0, T/3)$.

Pour construire ce troisième schéma, on repart de l'expression intégrale de I :

$$\begin{aligned}
I &= \frac{S_0}{T}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \exp\left((r-\frac{\sigma^2}{2})u + \sigma W_u\right) du \\
&= \frac{1}{T}\sum_{k=0}^{N-1} S_{t_k}\int_{t_k}^{t_{k+1}} \exp\left((r-\frac{\sigma^2}{2})(u-t_k) + \sigma(W_u - W_{t_k})\right) du \\
&\simeq \frac{1}{T}\sum_{k=0}^{N-1} S_{t_k}\left(\delta t + (r-\frac{\sigma^2}{2})\frac{(\delta t)^2}{2} + \sigma\int_{t_k}^{t_{k+1}} W_u\,du - \sigma\delta t W_{t_k}\right)
\end{aligned}$$

Ceci fournit donc la troisième approximation de l'intégrale I :

$$I_3 = \frac{1}{T}\sum_{k=0}^{N-1} S_{t_k}\left(\delta t + (r-\frac{\sigma^2}{2})\frac{(\delta t)^2}{2} + \sigma\int_{t_k}^{t_{k+1}} W_u\,du - \sigma\delta t W_{t_k}\right).$$

	$M = 10000$	$M = 100000$	$M = 1000000$
Schéma 1	7.2705 [7.0414, 7.4996]	7.3266 [7.2527, 7.4006]	7.2550 [7.2318, 7.2782]
Schéma 2	7.3269 [7.0960, 7.5578]	7.3831 [7.3086, 7.4576]	7.3112 [7.2878, 7.3346]
Schéma 3	7.4075 [7.1733, 7.6417]	7.3271 [7.2531, 7.4011]	7.3065 [7.2831, 7.3298]

FIG. 1 – Schémas de base : valeur estimée [intervalle de confiance] convergence de Monte Carlo ($N = 100$)

	$N = 10$	$N = 100$	$N = 1000$
Schéma 1	6.6673 [6.5992, 6.7354]	7.2274 [7.1541, 7.3008]	7.2942 [7.2200, 7.3685]
Schéma 2	7.2143 [7.1406, 7.2880]	7.2834 [7.2095, 7.3573]	7.2999 [7.2256, 7.3742]
Schéma 3	7.3073 [7.2332, 7.3814]	7.2869 [7.2129, 7.3608]	7.2508 [7.1768, 7.3248]

FIG. 2 – Schémas de base : valeur estimée [intervalle de confiance] convergence en temps ($M = 100000$)

En pratique, connaissant W_{t_k}, on simule d'abord $W_{t_{k+1}}$ puis, en utilisant le lemme ci-dessus, on simule $\int_{t_k}^{t_{k+1}} W_u\,du$ sachant W_{t_k} et $W_{t_{k+1}}$. La procédure correspondant à ce schéma est la procédure Schema3. Les résultats sont donnés dans les tableaux 1 et 2.

Remarque : On montre que les approximations I_1 et I_2 convergent en norme L^2 vers I en $O\left(\frac{1}{N}\right)$ alors que I_3 converge en norme L^2 vers I en $O\left(\frac{1}{N\sqrt{N}}\right)$ (cf. [LT]).

2.2 Méthodes de réduction de variance par variables de contrôle

Pour accélérer la convergence des méthodes de Monte Carlo, on peut utiliser des méthodes de réduction de variance, qui consistent à réduire la variance de la variable aléatoire simulée par Monte Carlo, de manière à améliorer la convergence vers l'espérance. Rappelons une conséquence du théorème central limite : si $(X_i)_{1\le i\le n}$ désigne n réalisations d'une même variable aléatoire X dont on cherche à calculer l'espérance, on a le résultat suivant :

$$P\left(\frac{1}{M}\sum_{i=1}^{M} X_i \in \left[\mathbb{E}(X) - 1.96\frac{\theta}{\sqrt{M}}, \mathbb{E}(X) + 1.96\frac{\theta}{\sqrt{M}}\right]\right) \approx 0.95$$

où $\theta^2 = \mathbb{E}((X - \mathbb{E}(X))^2)$ désigne la variance de X). La convergence est donc en $O\left(\frac{\theta}{\sqrt{M}}\right)$. Une première méthode pour réduire la variance consiste à écrire $\mathbb{E}(X) = \mathbb{E}(X-Y) + \mathbb{E}(Y)$ avec Y une variable aléatoire dont on connaît l'espérance, et qui est suffisemment proche de la variable aléatoire X pour que $\mathrm{Var}(X - Y) < \mathrm{Var}(X)$. Cette méthode s'appelle la réduction de variance par *variable de contrôle*.

On cherche par exemple à simuler des calls (ce qui suit s'adapte sans difficulté à des puts).

On utilise ici une idée de Kemma et Vorst (cf. les références dans [LT]) qui consiste à écrire $\frac{1}{T}\int_0^T S_u\,du \simeq \exp\left(\frac{1}{T}\int_0^T \ln(S_u)\,du\right)$ (ce qui n'est pas trop faux si r et σ ne sont pas trop grands). On considère donc la variable aléatoire $I' = \frac{1}{T}\int_0^T \ln(S_u)\,du = \ln(S_0) + (r - \frac{\sigma^2}{2})\frac{T}{2} + \sigma\frac{1}{T}\int_0^T W_t\,dt$ qui suit la loi $\mathcal{N}\left(\ln(S_0) + \left(r - \frac{\sigma^2}{2}\right)\frac{T}{2}, \frac{\sigma^2 T}{3}\right)$ (cf. le lemme 3) et on

	$M = 10000$	$M = 100000$	$M = 1000000$
Reduc 1	7.3119 [7.2974, 7.3264]	7.3173 [7.3125, 7.3221]	7.3144 [7.3129, 7.3159]
Reduc 2	7.3056 [7.2912, 7.3200]	7.3109 [7.3062, 7.3157]	7.3081 [7.3066, 7.3096]
Reduc 3	7.3125 [7.2974, 7.3277]	7.3092 [7.3044, 7.3140]	7.3067 [7.3052, 7.3082]

FIG. 3 – Schémas avec variable de contrôle : valeur estimée [intervalle de confiance] convergence de Monte Carlo ($N = 100$)

connaît donc explicitement $e^{-rT}\mathbb{E}((\exp(I') - K)^+)$ (c'est la formule de Black et Scholes). On rappelle le résultat dans le lemme suivant :

Lemme 4 *Soit G une variable aléatoire gaussienne de moyenne m et de variance s. Soit K un réel. Alors,*

$$\mathbb{E}\left((e^G - K)^+\right) = e^{m+s/2} N\left(\sqrt{s} - \frac{\ln(Ke^{-m})}{\sqrt{s}}\right) - KN\left(-\frac{\ln(Ke^{-m})}{\sqrt{s}}\right)$$

où $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$.

Preuve : Faire le calcul... $\diamondsuit$

Remarque : La fonction N est contenue dans la plupart des langages de programmation (ou plus souvent la fonction erreur erf avec $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$ et $N(x) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) + 1\right)$).

Nous écrivons donc, pour implémenter la réduction de variance :

$$\mathbb{E}\left((I - K)^+\right) = E\left((I - K)^+ - (\exp(I') - K)^+\right) + E\left((\exp(I') - K)^+\right).$$

La variable de contrôle que l'on utilise est donc (cf. I') $Z = e^{-rT}(\exp(I') - K)^+ = e^{-rT}\left(S_0 \exp\left((r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma}{T}\int_0^T W_u\, du\right) - K\right)^+$. Pour chacun des schémas précédents, on utilise le même type de discrétisation pour approcher $\frac{1}{T}\int_0^T W_u\, du$ dans la variable aléatoire $(I - K)^+$ et dans Z, en utilisant les mêmes séries de trajectoire. On détaille quelles sont exactement les expressions utilisées dans chacun des schémas précédents.

Pour le premier schéma, on utilise dans la variable de contrôle $\frac{1}{T}\int_0^T W_u\, du = \frac{1}{N}\sum_{k=0}^{N-1} W_{t_k}$.

Pour le deuxième schéma, on utilise dans la variable de contrôle $\frac{1}{T}\int_0^T W_u\, du = \frac{1}{N}\sum_{k=0}^{N-1} \frac{W_{t_k} + W_{t_{k+1}}}{2}$. On utilise donc exactement la méthodes des trapèzes.

Pour le troisième schéma, on utilise dans la variable de contrôle $\frac{1}{T}\int_0^T W_u\, du = \frac{1}{T}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} W_u\, du$. En effet, chacune de ces quantités est simulée directement dans le schéma.

Les procédures correspondantes dans le programme sont SchemaRed1, SchemaRed2 et SchemaRed3. On donne les résultats dans les tableaux 3 et 4.

2.3 Méthodes de réduction de variance par échantillonnage préférentiel

Une autre idée pour réduire la variance est de biaiser la variable simulée, puis ensuite de corriger ce biais. La remarque importante est la suivante : quand on simule un call, les réalisations qui ont une influence sur le résultat sont celles qui sont supérieures au strike K (car sinon, $\phi(I) = 0$) ; le résultat est donc d'autant moins précis que la plupart des

	$N=10$	$N=100$	$N=1000$
Reduc 1	7.3620 [7.3569, 7.3670]	7.3124 [7.3076, 7.3172]	7.3091 [7.3044, 7.3139]
Reduc 2	7.3050 [7.3004, 7.3096]	7.3062 [7.3014, 7.3110]	7.3085 [7.3037, 7.3132]
Reduc 3	7.3048 [7.3001, 7.3095]	7.3039 [7.2992, 7.3087]	7.3058 [7.3011, 7.3106]

FIG. 4 – Schémas avec variables de contrôle : valeur estimée [intervalle de confiance] convergence en temps ($M = 100000$)

réalisations sont plus petites que K et que très peu sont plus grandes. L'idée est alors de décaler les réalisations pour qu'elles soient assez souvent au-dessus de K et de corriger ensuite ce biais. C'est une méthode *d'échantillonnage préférentiel* (importance sampling en anglais). Evidemment, cette méthode n'a d'intérêt que si $S_0 < K$ (dans le cas d'un call). Dans la suite je n'ai essayé d'appliquer les méthodes de réduction de variance par importance sampling que sur le schéma de base 1.

Parité call put

Suite à cette remarque, on peut s'apercevoir qu'il est plus précis de simuler un put qu'un call, car les valeurs prises par un put sont bornées. Ceci peut constituer une méthode efficace de réduction de variance par variable de contrôle : on contrôle le call par un put. Il faut pour cela disposer d'une formule de parité call-put, que l'on a bien dans le cas des options asiatiques.

Lemme 5

$$\mathbb{E}\left(\frac{1}{T}\int_0^T S_u\,du - K\right) = \frac{S_0}{rT}(e^{rT}-1) - K$$

On utilise ce lemme (obtenu par un simple calcul) en écrivant que $\mathbb{E}((I-K)^+) = \mathbb{E}(I-K) + \mathbb{E}((K-I)^+)$: il suffit donc de simuler le put pour avoir la valeur du call.

Remarque : Cette méthode correspond plutôt à une méthode de variable de contrôle mais je l'ai incluse ici car l'idée de base qui permet de justifier qu'il vaut mieux simuler un put qu'un call ressemble à l'idée de base de l'importance sampling.

L'échantillonnage préférentiel

En pratique, voici comment se programme l'échantillonnage préférentiel. On veut calculer $\mathbb{E}(f(X))$ avec X une variable aléatoire de densité de probabilité $p_X(x)$. Au lieu de simuler X, on simule $\tilde{X}$, de densité de probabilité $p_{\tilde{X}}(x)$ et on calcule alors $\tilde{E}(f(\tilde{X})f_c(\tilde{X}))$ avec $f_c(x) = \frac{p_X(x)}{p_{\tilde{X}}(x)}$ le facteur de correction. Il est clair qu'avec ce facteur de correction, on a $\mathbb{E}(f(X)) = \tilde{E}(f(\tilde{X})f_c(\tilde{X}))$.

Première méthode

Une première idée pour décaler les trajectoires est de faire partir le brownien de $m > 0$ au lieu de 0. Disons-le tout de suite : ça ne marche pas très bien...

Dans le cadre du schéma 1, on note donc $X = (W_{t_1}, W_{t_2}, ..., W_{t_{N-1}})$ la variable aléatoire simulée et $\tilde{X} = X + \overline{m}$ avec $\overline{m} = (m, m, ..., m)$. X est un vecteur gaussien dont on peut facilement trouver la matrice de variance covariance, en remarquant que $X = AZ$ avec $A_{i,j} = 1$ si $i \leq j$ et 0 sinon et $Z = (W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_{N-1}} - W_{t_{N-2}})$. La matrice de variance covariance de Z est $\delta t Id$ donc celle de X est $\Gamma = \delta t A({}^t A)$. On a donc la densité de probabilité de X par le lemme suivant.

Lemme 6 *Soit X un vecteur gaussien à valeurs dans $\mathbb{R}^d$ centré de matrice de variance covariance Γ. Alors la densité de probabilité de X est*

$$P(X) = \frac{1}{\sqrt{2\pi}^d \sqrt{\det(\Gamma)}} \exp\left(-\frac{{}^tX\Gamma^{-1}X}{2}\right)$$

Preuve : Ce résultat est clairement vrai si Γ est une matrice diagonale, car dans ce cas, les coordonnées de X suivent des lois normales indépendantes centrées de variance le terme diagonale correspondant.

On diagnonalise la matrice symétrique Γ dans une base orthonormée : $\Gamma =^t \Omega D\Omega$ avec Ω matrice orthogonale et D diagonale de diagonale $(\lambda_1, ..., \lambda_d)$. Posons $Y = \Omega X$. Y est également un vecteur gaussien, de matrice de variance covariance $\mathbb{E}(Y \otimes Y) = \mathbb{E}(\Omega X \otimes \Omega X) = \Omega\Gamma^t\Omega = D$. Donc on a :

$$\begin{aligned}
\mathbb{E}(f(Y)) &= \int f(y)\frac{1}{\sqrt{2\pi}^d \sqrt{\det(D)}} \exp\left(-\frac{{}^tyD^{-1}y}{2}\right) dy \\
&= \int f(x)\frac{1}{\sqrt{2\pi}^d \sqrt{\det(\Gamma)}} \exp\left(-\frac{{}^tx^t\Omega D^{-1}\Omega x}{2}\right) |\det(\Omega)|\, dx \\
&= \int f(x)\frac{1}{\sqrt{2\pi}^d \sqrt{\det(\Gamma)}} \exp\left(-\frac{{}^tx\Gamma^{-1}x}{2}\right) dx
\end{aligned}$$

$\diamondsuit$

Appliquer la réduction de variance revient donc à simuler $\mathbb{E}\left(\phi(X+\overline{m})\frac{p_X(X+\overline{m})}{p_X(X)}\right)$. Quand on fait les calculs, on trouve que le facteur de correction ne dépend en fait que de m et de la première coordonnée de X : W_{t_1}. Le problème est de choisir m, et je n'ai pas réussi à trouver de formule très efficace (contrairement à la méthode qui suit). La programmation de cette méthode est la procédure SchemaSampling1.

Je pense que si les résultats ne sont pas très bons, c'est parce que le facteur de correction peut être très grand, ou encore, que les deux variables aléatoires X et $\tilde{X}$ sont trop différentes (pour appliquer la méthode, il faut en effet que les probabilités de X et $\tilde{X}$ soient équivalentes, et elles sont presque étrangères dans notre cas.)

Deuxième méthode

La deuxième méthode repose sur le lemme de Girsanov. Une manière plus naturelle de décaler la trajectoire est de remplacer W_t par $\tilde{W}_t = W_t + \theta t$ on sait alors que si $d\tilde{P} = e^{-\theta W_T - \frac{\theta^2}{2}T}dP$, alors $\tilde{E}(f(\tilde{W}_t, 0 \le t \le T)) = E(f(W_t, 0 \le t \le T))$ avec $\tilde{E}$ l'espérance sous la probabilité $\tilde{P}$ (f peut dépendre de toute la trajectoire de W_t). On a donc

$$\mathbb{E}(e^{-\theta W_T - \frac{\theta^2}{2}T} f(W_t + \theta t, 0 \le t \le T)) = \mathbb{E}(f(W_t, 0 \le t \le T)).$$

On obtient donc ainsi une méthode d'importance sampling, programmée dans la procédure SchemaSampling2.

Une bonne méthode pour trouver un θ approprié est de chercher à emmener $\tilde{I}$ autour de K. Concrètement, cela revient à résoudre $\frac{1}{T}\int_0^T S_0 \exp((r - \frac{\sigma^2}{2})t + \sigma\theta t) = K$.

	$M = 10000$	$M = 100000$	$M = 1000000$
Schéma de base 1	0.3493 [0.3055, 0.3931]	0.3881 [0.3731, 0.4031]	0.3760 [0.3712, 0.3807]
Sampling 1	0.3274 [0.2856, 0.3693]	0.3873 [0.3722, 0.4023]	0.3712 [0.3666, 0.3758]
Sampling 2	0.3643 [0.3479, 0.3806]	0.3765 [0.3707, 0.3822]	0.3757 [0.3739, 0.3775]

FIG. 5 – Schémas avec échantillonnage préférentiel : valeur estimée [intervalle de confiance] convergence de Monte Carlo $S_0 = 75$ $N = 100$, Sampling 1 : $m = 0.01$, Sampling 2 : $\theta = 2.305556$

3 Formulation sous forme d'Equation aux Dérivées Partielles

Une approche totalement différente est offerte grâce à la formulation du problème sous forme d'une EDP. On renvoie à [RS] pour les détails des calculs qui permettent d'obtenir l'EDP.

On considère la fonction $\phi(t,x) = \mathbb{E}\left((\frac{1}{T}\int_t^T S_u\,du - x)^+ | S_t = 1\right)$ et $f(t,x) = e^{-r(T-t)}\phi(t,x)$. Alors, le prix de l'option asiatique de strike K est donné par $P = S_0 f(0, \frac{K}{S_0})$ et f vérifie l'EDP :

$$\left\{ \begin{array}{rcl} \frac{\partial f}{\partial t} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 f}{\partial x^2} - (\frac{1}{T} + rx)\frac{\partial f}{\partial x} & = & 0 \\ f(T,x) & = & x^- \end{array} \right.$$

Par ailleurs, on montre que pour $x \le 0$, $f(t,x) = \frac{1}{rT}(1 - e^{-r(T-t)}) - xe^{-r(T-t)}$ et que quand $x \longrightarrow \infty$, $f(t,x) = 0$ (cf. la définition de f), ce qui donne des conditions aux limites.

Pour la discrétisation de cette EDP, on remarque que l'on a tout d'abord un opérateur type Black Scholes $L_1 = \frac{\sigma^2 x^2}{2}\frac{\partial^2}{\partial x^2} - rx\frac{\partial}{\partial x}$ et un opérateur d'advection $L_2 = -\frac{1}{T}\frac{\partial}{\partial x}$. Le problème est que l'advection prend le pas sur la diffusion quand x est petit, et c'est pour x petit qu'on a besoin de précision (cf. l'expression de P).
Pour la discrétisation, j'ai utilisé en général des schémas en temps type Cranck- Nicholson, un schéma de discrétisation de l'opérateur L_1 classique de Black et Scholes, et diverses discrétisations pour L_2 (cf. [BP]). Pour réaliser la discrétisation de L_1, on met l'équation vérifiée par f sous la forme :

$$\frac{\partial f}{\partial t} + \frac{\sigma^2}{2}\frac{\partial}{\partial x}\left(x^2\frac{\partial f}{\partial x}\right) - (r+\sigma^2)x\frac{\partial f}{\partial x} - \frac{1}{T}\frac{\partial f}{\partial x} = 0$$

Je donne ci-dessous une revue de ce que j'ai essayé : ce qui marche le mieux est un des schémas les plus simples : un schéma centré sur L_2.

- *Diff2.m* : c'est le schéma le plus simple : pour discrétiser L_2, on utilise au choix un schéma centré, un schéma type Lax-Wendroff ou un schéma centré/décentré (upwind) : plus la diffusion est faible, plus il est centré et inversement. (cf. chapitre 14 de [PR]).
- *Diff4.m* : c'est un schéma de splitting entre les deux opérateurs L_1 et L_2 : à chaque pas de temps, on résoud d'abord Black et Scholes (L_1), puis on transporte la solution (L_2) en utilisant une méthode de type volume fini.
- *Diff6.m* : on fait le changement de variable $g(t,x) = f(t, x - t/T)$, ce qui permet d'éliminer L_2 et on utlise l'expression explicite de f pour $x < 0$ pour donner des conditions aux limites sur g.

	valeur exacte	centré	décentré	mixte
$\sigma = 0.3\, I = 100\, N = 1000$	7.308	7.295	9.299	8.217
$\sigma = 0.3\, I = 600\, N = 1000$	7.308	7.3043	8.3667	7.6574
$\sigma = 0.3\, I = 600\, N = 500$	7.308	7.3041	8.3666	7.6572
$\sigma = 0.3\, I = 1000\, N = 200$	7.308	7.3059	7.4756	7.9606
$\sigma = 0.05\, I = 600\, N = 500$	1.697	1.783	4.612	4.453

FIG. 6 – Discrétisation de l'EDP selon trois schémas pour L_2 : centré, décentré ou mixte centré/décentré. $S_0 = 100$ $K = 100$ $r = 0.02$

L'équation vérifiée par g est :

$$\frac{\partial g}{\partial t} + \frac{\sigma^2}{2}\frac{\partial}{\partial x}\left((x - \frac{t}{T})^2 \frac{\partial g}{\partial x}\right) - (r + \sigma^2)(x - \frac{t}{T})\frac{\partial g}{\partial x} = 0$$

et on a : $g(x,t) = (x-1)^-$ et $g(0,t) = f(-\frac{t}{T}, t)$.

– *Diff7.m* : on fait le changement de variable $h(t,x) = f(t, e^x)$ pour simplifier L_1.
L'équation vérifiée par h est :

$$\frac{\partial h}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 h}{\partial x^2} - (r + \frac{\sigma^2}{2})\frac{\partial h}{\partial x} - \frac{1}{T}e^{-x}\frac{\partial h}{\partial x} = 0$$

De toute ces méthodes, c'est encore la première qui marche le mieux. Plus σ est petit, plus les résultats sont mauvais... On donne quelques résultats obtenus avec Diff2. Numériquement, voici les paramètres utilisés : $0 \le x \le x_{max}$ avec $x_{max} = 6$, on note N le paramètre de discrétisation en temps (le pas de temps est T/N) et I le paramètre de discrétisation en espace (le pas d'espace est x_{max}/I). J'ai choisi x_{max} en calculant la moyenne et la variance de $\frac{1}{T}\int_0^T S_u\, du$.

4 Conclusion

En conclusion, on peut essayer de dresser un bilan de la comparaison des deux approches.
L'approche par l'EDP a l'avantage de ne pas dépendre en précision de S_0 et K_0 mais elle souffre d'une trop mauvaise précision, surtout quand σ devient petit. Je n'ai pas réussi à trouver une bonne discrétisation pour lever cette imprécision.
L'approche Monte Carlo a l'avantage de se programmer très facilement (quand on a les équations, on a le programme), d'être assez précise et de ne pas dépendre de la valeur de σ (et même, plus σ est petit, plus l'intervalle de confiance se resserre !). De plus, elle a l'avantage de donner un intervalle de confiance. Par contre, la méthode dépend de S_0 et K_0 (bien qu' on arrive à rester précis grâce à des méthodes d'importance sampling).
Sur ces équations, les approches numériques probabilistes sont donc apparemment plus compétitives.

Références

[BP] H. BERESTYCKI O. PIRONNEAU, *Introduction aux méthodes mathématiques en finance*, cours DEA, février-juin 2000.

[LL] D. LAMBERTON B. LAPEYRE, *Introduction au calcul stochastique appliqué à la finance*, ellipses

[LT] B. Lapeyre E. Temam, *Competitive Monte Carlo methods for the Pricing of Asian Options*, preprint (février 2000).

[PR] M. Picasso J. Rappaz,*Introduction à l'analyse numérique*, Presses polytechniques et universitaires romandes, (1998).

[RS] L.C.G. Rogers Z. Shi, *The value of an Asian option*, J. Appl. Probability, 32, p 1077-1088, (1995).