Méthodes Monte Carlo en Finance

Chapitre 5-1
Techniques de réduction de la variance : variable antithétique, variable de contrôle et appariement des moments

OUESLATI Amor

Université de Manouba
Ecole Nationale des Sciences de L'Informatique
Département des Méthodes Quantitatives

11 janvier 2014

Estimateur de Monte Carlo

- Soit $X_1, X_2, ..., X_n$ des variables aléatoires indépendantes et identiquement distribuées. Elles sont obtenues à l'aide du générateur de nombre aléatoire de votre logiciel préféré.
- L'estimateur de Monte Carlode la quantité

$$\theta \equiv E[g(X)]$$

est

$$\hat{\theta}_n \equiv \frac{1}{n} \sum_{i=1}^{n} g(X_i).$$

- Remarquons que $\hat{\theta}_n$ est une fonction de l'échantillon $X_1, X_2, ..., X_n$. C'est une variable aléatoire.

Notation

1. Le paramètre devant être estimé : $\theta = E[g(X)]$.
2. La variance :$\sigma^2 = Var[g(X)] < \infty$.
3. L'estimateur : $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i)$
4. La variance échantillonnale :

$$\left(\widehat{\sigma}^{(n)}\right)^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(g(X_i) - \hat{\theta}_n\right)^2$$

5. z_α est le quantile d'ordre $1-\alpha$ de la distribution normale centrée et réduite, c'est-à-dire que

$$Pr[Z \leq z_\alpha] = 1 - \alpha.$$

Propriétés statistiques

1. $\hat{\theta}_n$ est un estimateur sans biais et convergent pour θ.
2. $\hat{\theta}_n$ converge presque sûrement vers la quantité à estimer θ lorsque la taille nde l'échantillon tend vers l'infini.
3. Le théorème limite central implique que

$$\sqrt{n}\frac{\hat{\theta}_n - \theta}{\sigma} \overset{n \to \infty}{\Rightarrow} N(0,1).$$

4. L'intervalle de confiance pour θ (de niveau de confiance $1-\alpha$) est $\left[\hat{\theta}_n - z_{\frac{\alpha}{2}}\widehat{\sigma}^{(n)}\sqrt{n}; \hat{\theta}_n + z_{\frac{\alpha}{2}}\widehat{\sigma}^{(n)}\sqrt{n}\right]$.
5. La marge d'erreur est $z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \cong z_{\frac{\alpha}{2}}\widehat{\sigma}^{(n)}\sqrt{n}$.

Marge d'erreur

- La marge d'erreur est $z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}} = z_{\frac{\alpha}{2}} \sqrt{Var\left[\hat{\theta}_n\right]}$
- Si on ne veut pas augmenter la taille de l'échantillon ni changer le niveau de confiance $1-\alpha$, il nous faut réduire la variance σ^2/n de l'estimateur.
- Est-il possible d'estimer θ à l'aide d'un nouvel estimateur dont la variance est plus petite que $Var\left[\hat{\theta}_n\right] = \sigma^2/n$?

Souvent, nous cherchons à estimer

$$\theta = E[g(Z)] = \int_{-\infty}^{\infty} g(z) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

où Z est une variable aléatoire gaussienne centrée et réduite.

L'idée de la variable antithétique est de profiter de la symétrie de la distribution de Z. Notons que

$$\theta = E[g(-Z)].$$

En effet,

$$\begin{aligned} E[g(-Z)] &= \int_{-\infty}^{\infty} g(-z)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= \int_{-\infty}^{\infty} g(w)\frac{1}{\sqrt{2\pi}} exp\left(-\frac{w^2}{2}\right) dw \text{ si } w = -z \\ &= E[g(Z)]. \end{aligned}$$

- Parce que l'espérance est un opérateur linéaire,

$$\theta = E\left[\frac{g(Z)+g(-Z)}{2}\right] = E[h(Z)]$$

où $h(z) = (g(z) + g(-z))/2$.

- Or,

$$\begin{aligned} \zeta^2 = Var[h(Z)] &= Var\left[\frac{g(Z)+g(-Z)}{2}\right] \\ &= \frac{1}{4}(Var[g(Z)] + Var[g(-Z)] + 2Cov[g(Z), g(-Z)]) \\ &= \frac{1}{2}(\sigma^2 + Cov[g(Z), g(-Z)]). \end{aligned}$$

- En créant un estimateur basé sur $h(Z)$ au lieu de $g(Z)$, nous aurons une réduction de la variance en autant que $Cov[g(Z), g(-Z)] < \sigma^2$.

- Notons que puisque la corrélation est bornée par 1, alors

$$
\begin{aligned}
1 &\geq Corr[g(Z), g(-Z)] \\
&= \frac{Cov[g(Z), g(-Z)]}{\sqrt{Var[g(Z)]}\sqrt{Var[g(-Z)]}} \\
&= \frac{Cov[g(Z), g(-Z)]}{\sigma^2}.
\end{aligned}
$$

- Ainsi,

$$Cov[g(Z), g(-Z)] \leq \sigma^2,$$

c'est-à-dire que l'usage d'une variable antithétique ne peut pas augmenter la variance.

- Le gain sera énorme si la fonction g est à peu près linéaire. En effet, le cas $g(x) = ax + b$ nous donne

$$\begin{aligned}
\xi^2 &= Var\left[\frac{g(Z) + g(-Z)}{2}\right] \\
&= Var\left[\frac{(aZ + b) + (-aZ + b)}{2}\right] \\
&= Var[b] \\
&= 0.
\end{aligned}$$

- D'autre part, il y a des fonctions pour lesquelles la variable antithétique ne fonctionne pas du tout : si $g(x) = ax^2$, alors

$$\begin{aligned}\xi^2 &= Var\left[\frac{g(Z)+g(-Z)}{2}\right]\\ &= Var\left[\frac{aZ^2+a(-Z)^2}{2}\right]\\ &= Var\left[aZ^2\right]\\ &= \sigma^2.\end{aligned}$$

L'estimateur de Monte Carlo de θ avec l'utilisation d'une variable antithétique est

$$\hat{\theta}_n^A \equiv \frac{1}{n}\sum_{i=1}^{n} h(Z_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \frac{g(Z_i) + g(-Z_i)}{2}$$

où $Z_1, Z_2, ..., Z_n$ sont des variables aléatoires indépendantes et identiquement distribuées de loi normale centrée et réduite.

- Puisque $\hat{\theta}_n^A \equiv n^{-1} \sum_{i=1}^{n} h(Z_i)$ est une somme de variables aléatoires indépendantes et identiquement distribuée, ses propriétés asymptotiques sont les mêmes que celles de l'estimateur de Monte Carlo en substituant σ^2 par ξ^2 :
- Sans biais

$$E\left[\hat{\theta}_n^A\right] = \frac{1}{n} \sum_{i=1}^{n} E[h(Z_i)] = \frac{1}{n} \sum_{i=1}^{n} \theta = \theta.$$

- Convergent

$$Var\left[\hat{\theta}_n^A\right] = \frac{1}{n^2} \sum_{i=1}^{n} Var[h(Z_i)] = \frac{\xi^2}{n}$$

$$\rightarrow 0 \text{ lorsque } n \rightarrow \infty.$$

- La variable aléatoire $\hat{\theta}_n^A$ converge presque sûrement vers θ lorsque $n \to \infty$.
- Le théorème limite central implique que

$$\sqrt{n}\frac{\hat{\theta}_n^A - \theta}{\xi}$$

converge en loi vers une distribution normale centrée et réduite lorsque la taille n de l'échantillon tend vers l'infini où

$$\xi^2 = \sigma^2 - \frac{\sigma^2 - Cov[g(Z), g(-Z)]}{2} \leq \sigma^2.$$

- $g(Z_i) = e^{-rT} \max \left\{ S_0 e^{\left(r - \frac{\sigma^2}{2}\right) T + \sigma \sqrt{T} Z_i} - K ; 0 \right\}$
- $h(Z_i) = \dfrac{g(Z_i) + g(-Z_i)}{2}$

Z_i	S_i	$g(Z_i)$	$S_i(-Z_i)$	$g(-Z_i)$	$h(Z_i)$
0.709	89.42	0	118.75	27.35	13.67
0.044	103.96	13.28	102.14	11.54	12.41
0.241	98.20	7.80	108.12	17.24	12.52
0.973	84.81	0	125.19	33.47	16.74

$S_0 = 100$, $r = 5\%$, $\nu = 0.2, T = 1$, $K/S_0 = 0.9$, $\theta = 16.70$.

- Le prix S_t de l'actif sous-jacent à la $t^{\text{ième}}$ journée, sous la mesure neutre au risque Q, est, $\forall t \in \{0, 1, 2, ...\}$,

$$\ln \frac{S_{t+1}}{S_t} = \rho - \frac{1}{2} h_{t+1} + (h_{t+1})^{1/2} \epsilon_{t+1}$$
$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t (\epsilon_t - \widetilde{\lambda})^2$$

  où $\{\epsilon_t : t \in \mathbb{N}\}$ est une suite de variables aléatoires gaussiennes (d'espérance nulle et de variance unitaire) indépendantes sous la mesure Q.

- Le processus h représente la variance conditionnelle du rendement quotidien du titre : $h_{t+1} = Var_t^Q \left[\ln \frac{S_{t+1}}{S_t} \right]$.

- Les paramètres sont choisis de sorte que ce modèle de marché ait un comportement similaire à celui de Black et Scholes utilisé dans l'exemple.
  - $S_0 = 100$.
  - Le taux d'intérêt sans risque ρ est quotidien. Ainsi $\rho \times 365 = r$, d'où $\rho = 0.05/365 = 1.3699 \times 10^{-4}$.
  - Nous prendrons des valeurs typiquement observées pour les paramètres β : $\beta_0 = 0.00001$, $\beta_1 = 0.8$ et $\beta_2 = 0.1$.
  - La variance stationnaire est $h^* = \beta_0 \left\{1 - \beta_1 - \beta_2[1 + \widetilde{\lambda}^2]\right\}^{-1}$. Pour que la variance stationnaire (quotidienne) soit équivalente à celle utilisée dans le cadre de Black et Scholes, nous posons $h^* = \beta_0 \left\{1 - \beta_1 - \beta_2[1 + \widetilde{\lambda}^2]\right\}^{-1} = \nu^2/365$ ce qui implique que $\widetilde{\lambda} = 0.29580$.
  - Finalement, nous posons $h_1 = h^* = 1.0959 \times 10^{-4}$.

Répondre aux questions suivantes en incluant les variables antithétiques. Il faut utiliser les mêmes valeurs de $\{\epsilon_t : t \in \mathbb{N}\}$ afin que les différences observées proviennent uniquement de l'utilisation des variables antithétiques et non pas d'un changement de nombres pseudo-aléatoires.

1. Déterminer le prix des options d'achat d'une maturité d'un an et dont les prix d'exercice sont respectivement $0.9 \times S_0$, S_0 et $1.1 \times S_0$ à l'aide d'une simulation de Monte Carlo pour des tailles d'échantillon égales à 5000, 10000, 50 000, 100 000, 500 000.et 1 000 000.
2. Construire les intervalles de confiance autour de ces prix.
3. Noter les temps de calcul nécessaires pour chaque taille d'échantillon.
4. Produisez un graphe présentant la marge d'erreur en abscisse et le temps de calcul en ordonnée.

5. Comparer vos résultats avec ceux obtenus sans réduction de la variance (présenter les deux courbes sur le même graphe) et commentez.

- Rappelons que le but est d'estimer $\theta = E[g(X)]$ où X est une variable aléatoire de distribution quelconque dont les deux premiers moments existent.
- Supposons maintenant qu'il existe une fonction $\gamma : \mathbb{R} \to \mathbb{R}$ telle que nous connaissons la quantité

$$\mu = E[\gamma(X)].$$

- Pouvons-nous utiliser cette information pour nous construire un estimateur dont la variance est plus petite que celle de l'estimateur de Monte Carlo ?

Puisque

$$\begin{aligned} E[g(X) + \beta(\gamma(X) - \mu)] &= E[g(X)] + \beta(E[\gamma(X)] - \mu) \\ &= \theta + \beta(\mu - \mu) \\ &= \theta \end{aligned}$$

où β est une constante, nous proposons un estimateur basé sur la variable aléatoire

$$h_\beta(X) = g(X) + \beta(\gamma(X) - \mu).$$

Notons que

$$
\begin{aligned}
& \mathit{Var}\left[h_\beta(X)\right] \\
& \qquad = \mathit{Var}[g(X) + \beta(\gamma(X) - \mu)] \\
& \qquad = \mathit{Var}[g(X)] + \mathit{Var}[\beta(\gamma(X) - \mu)] \\
& \qquad + 2\mathit{Cov}[g(X), \beta(\gamma(X) - \mu)] \\
& \qquad = \sigma^2 + \beta^2 \mathit{Var}[\gamma(X)] + 2\beta \mathit{Cov}[g(X), \gamma(X)].
\end{aligned}
$$

Historiquement, la constante β était supposée égale à 1. Il est tout de même possible de la choisir de sorte que $\mathit{Var}\left[h_\beta(X)\right]$ soit la plus petite possible.

Calculons la dérivée première :

$$\begin{aligned} &\frac{d}{d\beta} Var\left[h_\beta(X)\right] \\ &= \frac{d}{d\beta}\left(\sigma^2 + \beta^2 Var[\gamma(X)] + 2\beta Cov[g(X), \gamma(X)]\right) \\ &= 2\beta Var[\gamma(X)] + 2Cov[g(X), \gamma(X)]. \end{aligned}$$

Déterminons la valeur de β qui annule cette dérivée :

$$\widetilde{\beta} = \frac{Cov[g(X), \gamma(X)]}{Var[\gamma(X)]}.$$

Comme la dérivée seconde est positive

$$\left.\frac{d^2}{d\beta^2} Var\left[h_\beta(X)\right]\right|_{\beta=\widetilde{\beta}} = 2Var[\gamma(X)] > 0,$$

alors la variance minimale est atteinte lorsque $\beta = \widetilde{\beta}$.

Par conséquent, la variance minimale est

$$
\begin{aligned}
\xi^2_{\widetilde{\beta}} \equiv Var\left[h_{\widetilde{\beta}(X)}\right] & \\
&= \sigma^2 + \widetilde{\beta}^2 Var[\gamma(X)] + 2\widetilde{\beta} Cov[g(X), \gamma(X)] \\
&= \sigma^2 + \left(\frac{Cov[g(X), \gamma(X)]}{Var[\gamma(X)]}\right)^2 Var[\gamma(X)] \\
&- 2\frac{Cov[g(X), \gamma(X)]}{Var[\gamma(X)]} Cov[g(X), \gamma(X)] \\
&= \sigma^2 - \frac{(Cov[g(X), \gamma(X)])^2}{Var[\gamma(X)]} \\
&\leq \sigma^2.
\end{aligned}
$$

L'estimateur de Monte Carlo de θ avec variable de contrôle est

$$\begin{aligned}\widehat{\theta}_n^{\widetilde{\beta}} &\equiv \frac{1}{n}\sum_{i=1}^{n} h_{\widetilde{\beta}}(X_i)\\ &= \frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) + \widetilde{\beta}(\gamma(X_i) - \mu)\right]\end{aligned}$$

où $X_1, X_2, ..., X_n$ sont des variables aléatoires indépendantes et identiquement distribuées.

- Notre estimateur est une somme de variables aléatoires indépendantes et identiquement distribuées. Par conséquent, ses propriétés asymptotiques sont les mêmes que celles de l'estimateur de Monte Carlo en substituant σ^2 par $\xi^2_{\widetilde{\beta}}$.
- Sans biais :

$$E\left[\widehat{\theta}_n^{\widetilde{\beta}}\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} h_{\widetilde{\beta}}(X_i)\right] = E\left[h_{\widetilde{\beta}}(X)\right] = \theta.$$

- Convergent :

$$Var\left[\widehat{\theta}_n^{\widetilde{\beta}}\right] = \frac{\sigma^2}{n}\frac{(Cov[g(X),\gamma(X)])^2}{nVar[\gamma(X)]} \to 0 \text{ lorsque } n \to \infty.$$

- L'estimateur $\widehat{\theta}_n^{\widetilde{\beta}}$ converge presque sûrement vers θ lorsque $n \to \infty$.
- Le théorème limite central implique que

$$\sqrt{n}\frac{\widehat{\theta}_n^{\widetilde{\beta}} - \theta}{\xi_{\widetilde{\beta}}}$$

convergee n loi vers une distribution normale centrée et réduite lorsque la taille nde l'échantillon tend vers l'infini où

$$\xi_{\widetilde{\beta}}^2 = \sigma^2 - \frac{(Cov[g(X), \gamma(X)])^2}{Var[\gamma(X)]} \leq \sigma^2.$$

- En pratique, on ne peut déterminer la constante
$$\widetilde{\beta} = -\frac{Cov[g(X), \gamma(X)]}{Var[\gamma(X)]}$$
aisément.
- Le problème provient du fait que $Cov[g(X), \gamma(X)]$ et, parfois, $Var[\gamma(X)]$ sont des quantités inconnues.
- Rappelons que $\theta = E[g(X)]$ est la quantité que nous voulons estimer. Il est donc fort probable que $Cov[g(X), \gamma(X)] = E[g(X)\gamma(X)] - \theta E[\gamma(X)]$ soit aussi inconnue.

- Il est possible d'estimer

$$\widetilde{\beta} = -\frac{Cov[g(X), \gamma(X)]}{Var[\gamma(X)]}$$

  à l'aide de l'échantillon issu de la simulation de Monte Carlo. En effet, on peut utiliser les estimateurs de $Cov[g(X), \gamma(X)]$ et de $Var[\gamma(X)]$.

$$\widetilde{\beta}_n \equiv -\frac{\frac{1}{n}\sum_{i=1}^{n} g(X_i)\gamma(X_i) - \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right)\left(\frac{1}{n}\sum_{i=1}^{n} \gamma(X_i)\right)}{\frac{1}{n}\sum_{i=1}^{n}(\gamma(X_i))^2 - \left(\frac{1}{n}\sum_{i=1}^{n} \gamma(X_i)\right)^2}.$$

- La variable aléatoire $-\widetilde{\beta}_n$ est la pente de la droite des moindres carrés obtenue en régressant la variable $g(X)$ sur la variable $\gamma(X)$.

- $\widetilde{\beta}_n$ est une variable aléatoire qui dépend de tout l'échantillon (on devrait écrire $\widetilde{\beta}_n(X_1, \ldots X_n)$) contrairement à $\widetilde{\beta}$ qui est une constante (déterministe).
- Cela a pour conséquence de modifier le comportement de notre estimateur de Monte Carlo avec variable de contrôle. En effet,

$$\begin{aligned}\theta_n^{\widetilde{\beta}_n} &\equiv \frac{1}{n}\sum_{i=1}^{n} h_{\widetilde{\beta}(n)}(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) + \widetilde{\beta}^{(n)}(\gamma(X_i) - \mu)\right]\end{aligned}$$

n'est plus une somme de variables aléatoires indépendantes et identiquement distribuées.

- Mais son comportement asymptotique sera essentiellement le même :

$$\sqrt{n}\widehat{\theta}_n^{\widetilde{\beta}_n}$$
$$= \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) + \widetilde{\beta}_n(\gamma(X_i) - \mu)\right]$$
$$= \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) + \widetilde{\beta}(\gamma(X_i) - \mu)\right] + \frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{\beta}_n - \widetilde{\beta}\right)(\gamma(X_i) - \mu)\right]$$
$$= \underbrace{\sqrt{n}\widehat{\theta}_n^{\widetilde{\beta}}}_{\overset{n\to\infty}{\Rightarrow} N\left(\theta, \xi^2_{\widetilde{\beta}}\right) \text{ en loi}} + \underbrace{\left(\widetilde{\beta}_n - \widetilde{\beta}\right)}_{\overset{n\to\infty}{\to} p.s.} \underbrace{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\gamma(X_i) - \mu\right)}_{\overset{n\to\infty}{\Rightarrow} N(0, Var[\gamma(X)]) \text{ en loi}}.$$

Cependant, la variance en petit échantillon sera modifiée :

$$
\begin{aligned}
&Var\left[\widehat{\theta}_n^{\widetilde{\beta}_n}\right] \\
&= Var\left[\frac{1}{n}\sum_{i=1}^{n}\left[g(X_i)+\widetilde{\beta}_n(\gamma(X_i)-\mu)\right]\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n} Cov\left[\begin{array}{c} g(X_i)+\widetilde{\beta}_n(\gamma(X_i)-\mu), \\ g(X_j)+\widetilde{\beta}_n(\gamma(X_j)-\mu) \end{array}\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left(\begin{array}{c} cov\left[g(X_i),g(x_j)\right]+2Cov\left[g(X_i),\widetilde{\beta}_n\,(\gamma(X_j)-\mu)\right] \\ +Cov\left[\widetilde{\beta}_n(\gamma(X_i)-\mu),\widetilde{\beta}_n(\gamma(X_j)-\mu)\right] \end{array}\right) \\
&= Var\left[\widehat{\theta}_n^{\widetilde{\beta}}\right]\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left(\begin{array}{c} +2Cov\left[g(X_i),\widetilde{\beta}_n\,(\gamma(X_j)-\mu)\right] \\ +Cov\left[\widetilde{\beta}_n(\gamma(X_i)-\mu),\widetilde{\beta}_n(\gamma(X_j)-\mu)\right] \end{array}\right)
\end{aligned}
$$

Ledernier terme est le prix à payer pour avoir estimé $\widetilde{\beta}$.

- Il est possible d'utiliser plusieurs variables de contrôle en même temps.
- En effet, s'il existe des fonctions $\gamma_1, ..., \gamma_d$ telles que pour tout $j \in \{1, ..., d\}$, les quantités

$$\mu_j \equiv E[\gamma(X_j)]$$

sont connues, alors nous proposons un estimateur basé sur la variable aléatoire

$$h_\beta(X) = g(X) + \sum_{j=1}^{d} \beta_j \left(\gamma_j(X) - \mu_j\right).$$

- Sans biais pour θ :

$$E\left[h_\beta(X)\right] = \theta + \sum_{j=1}^{d} \beta_j \left(\mu_j - \mu_j\right) = \theta.$$

$$
\begin{aligned}
& Var\left[h_\beta(X)\right] \\
&= Var[g(X)] + 2\sum_{j=1}^{d} Cov\left[g(X), \beta_j\left(\gamma_j(X) - \mu_j\right)\right] \\
&+ \sum_{j=1}^{d}\sum_{j'=1}^{d} Cov\left[\beta_j\left(\gamma_j(X) - \mu_j\right), \beta_{j'}\left(\gamma_{j'}(X) - \mu_{j'}\right)\right] \\
&= Var[g(X)] + 2\sum_{j=1}^{d} \beta_j Cov\left[g(X), \gamma_j(X)\right] \\
&+ \sum_{j=1}^{d} \beta_j^2 Var\left[\gamma_j(X)\right] + 2\sum_{j=1}^{d}\sum_{j'=j+1}^{d} \beta_j\beta_{j'} Cov\left[\gamma_j(X), \gamma_{j'}(X)\right]
\end{aligned}
$$

Encore une fois, il est possible de choisir les constantes $\widetilde{\beta}_1, ..., \widetilde{\beta}_d$ de sorte que la variance soit minimale.

- En effet, si $\nabla Var\left[h_\beta(X)\right]$ dénote le vecteur de dérivées premières, alors $\widetilde{\beta}$, satisfaisant la relation $\nabla Var\left[h_{\widetilde{\beta}}(X)\right] = \overrightarrow{\mathbf{0}}$, rendra $Var\left[h_\beta(X)\right]$ optimale.
- Comme

$$\frac{\partial}{\partial \beta_i} Var\left[h_\beta(X)\right] = 2Cov[g(X), \gamma_i(X)] + 2\sum_{j=1}^{d} \beta_j Cov\left[\gamma_i(X), \gamma_j(X)\right],$$

alors le système d'équations linéaires à résoudre est :

$$\underbrace{\left(Cov\left[\gamma_i(X), \gamma_j(X)\right]\right)_{i,j=1,\dots,d}}_{d\times d} \underbrace{\widetilde{\beta}}_{d\times 1} = \underbrace{\left(Cov[g(X), \gamma_i(X)]\right)_{i=1,\dots d}}_{d\times 1}.$$

- Puisque $\left(Cov\left[\gamma_i(X),\gamma_j(X)\right]\right)_{i,j=1,\ldots,d}$ est une matrice de variances-covariances, elle est définie positive et est donc inversible. La solution est donc

$$\widetilde{\beta} = \underbrace{-\left(Cov\left[\gamma_i(X),\gamma_j(X)\right]\right)^{-1}_{i,j=1,\ldots,d}}_{d\times d}\underbrace{(Cov[g(X),\gamma_i(X)])_{i=1,\ldots d}}\, d\times 1.$$

- La matrice de dérivées secondes

$$\mathbf{H}_\beta = \left(\frac{\partial^2}{\partial\beta_i\partial\beta_j} Var\left[h_\beta(X)\right]\right)_{i,j=1,\ldots,d} = 2\underbrace{\left(Cov\left[\gamma_i(X),\gamma_j(X)\right]\right)_{i,j=1,\ldots,d}}_{d\times d}$$

est définie positive (c'est 2 fois une matrice de variances-covariances). Par conséquent, $\widetilde{\beta}$ fera en sorte que

$$Var\left[h_{\widetilde{\beta}}(X)\right] \leq Var\left[h_\beta(X)\right] \text{ pour tout } \beta\in\mathbb{R}^d.$$

L'estimateur de Monte Carlo de θ avec plusieurs variables de contrôle est

$$\widehat{\theta}_n^{\widetilde{\beta}} = \frac{1}{n}\sum_{i=1}^{n} h_{\widetilde{\beta}}(X_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) + \sum_{j=1}^{d}\widetilde{\beta}_j\left(\gamma_j(X_i) - \mu_j\right)\right]$$

où $X_1, X_2, ..., X_n$ sont des variables aléatoires indépendantes et identiquement distribuées.

Multiples variables de contrôle

efficience

- L'ajout de variables de contrôle réduit la variance mais augmente le temps de calcul à chaque trajectoire.
- Il y a donc un compromis à faire entre la réduction de la variance et le temps de calcul.
- Il vient un temps où l'ajout d'une variable de contrôle supplémentaire diminue l'efficacité (précision versus temps de calcul) de l'estimateur.

- $g(S_i) = \exp(-rT) \max \{S_i - K; 0\}$
- $\beta = -1$

Z_i	S_i	$S_i - \phi$	$g(S_i)$	$h(S_i)$
0.312	96.82	8.31	6.48	14.79
0.294	109.28	4.15	18.34	14.19
1.196	81.12	24.01	0	24.01
0.002	103.01	2.12	12.37	14.49

$S_0 = 100$, $r = 5\%$, $\sigma = 0.2$, $\phi = 105.13$, $\beta = -1$, $T = 1$, $K/S_0 = 0.9$, $\theta = 16.70$.

- $g(S_i) = \exp(-rT)\max\{S_i - K; 0\}$
- $\beta = -0.77$

Z_i	S_i	$S_i - \phi$	$g(S_i)$	$h(S_i)$
0.312	96.82	8.31	6.48	12.88
0.294	109.28	4.15	18.34	15.15
1.196	81.12	24.01	0	18.49
0.002	103.01	2.12	12.37	14.00

$S_0 = 100$, $r = 5\%$, $\sigma = 0.2$, $\phi = 105.13$, $\beta = -0.77$, $T = 1$, $K/S_0 = 0.9$, $\theta = 16.70$.

Reprendre la simulation de l'exercice précédent (dans le contexte NGARCH) en incluant comme variable de contrôle le prix de l'option dans le cadre du modèle de Black et Scholes.

1. Déterminer le prix des options d'achat d'une maturité d'un an et dont les prix d'exercice sont respectivement $0.9 \times S_0$, S_0 et $1.1 \times S_0$ à l'aide d'une simulation de Monte Carlo pour des tailles d'échantillon égales à 5000, 10000, 50 000, 100 000, 500 000.et 1 000 000.
2. Construire les intervalles de confiance autour de ces prix.
3. Noter les temps de calcul nécessaires pour chaque taille d'échantillon.
4. Produisez un graphe présentant la marge d'erreur en abscisse et le temps de calcul en ordonnée.

Travail pratique 2 II

5. Comparer vos résultats avec ceux obtenus sans réduction de la variance et avec l'utilisation de la variable antithétique (présenter les courbes sur le même graphe) et commentez.

Il faut utiliser les mêmes valeurs de $\{\epsilon_t : t \in \mathbb{N}\}$ afin que les différences observées proviennent uniquement de l'utilisation de la variable de contrôle et non pas d'un changement de nombres pseudo-aléatoires.

- Le but est encore d'estimer

$$\theta = E[g(X)]$$

  où X est une variable aléatoire de distribution quelconque dont les deux premiers moments existent.

- Dans bien des cas, nous connaissons l'espérance

$$\mu_X \equiv E[X]$$

  de la variable aléatoire X.

- Lors de la simulation des n copies indépendantes de X, $X_1, \dots, X_n$, il est fort probable que la moyenne échantillonnale soit différente de l'espérance théorique, c'est-à-dire que

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \cong \mu_X .$$

- L'idée derrière l'appariement du premier moment est de s'assurer que la moyenne échantillonnale soit égale à l'espérance théorique.

- Après avoir simuler les variables $X_1, \ldots, X_n$, nous en construisons de nouvelles à partir des premières :

$$X_i^{*(n)} = X_i - \overline{X}_n + \mu_X.$$

- Remarquons que

$$\overline{X}_n^{*(n)} = \frac{1}{n}\sum_{i=1}^{n} X_i^{*(n)} = \overline{X}_n - \overline{X}_n + \mu_X = \mu_X.$$

- De plus, $X_1^{*(n)}, \ldots, X_n^{*(n)}$ sont toujours identiquement distribuées mais elles ne sont plus indépendantes. Cela rendra l'obtention des propriétés asymptotiques de notre estimateur plus difficile.

L'estimateur proposé est

$$\widehat{\theta}_n^{M1} = \frac{1}{n} \sum_{i=1}^{n} g\left(X_i^{*(n)}\right).$$

- Si la fonction gest différentiable et que sa dérivée est continue, alors le développement en série de Taylor de g autour du point x_0 nous donne

$$g(x) = g(x_0) + g'(y)(x - x_0)$$

pour un certain y satisfaisant la contrainte

$$\min(x, x_0) \leq y \leq \max(x, x_0).$$

- Remplaçant x par $X_i^{*(n)} = X_i - \overline{X}_n + \mu_X$ et x_0 par X_i, nous obtenons

$$g\left(X_i^{*(n)}\right) = g(X_i) + g'\left(Y_i^{(n)}\right)\left(X_i^{*(n)} - X_i\right) \quad = g(X_i) - g'\left(Y_i^{(n)}\right)\left(\overline{X}_n - \mu_X\right)$$

où $Y_i^{(n)}$ est une variable aléatoire telle que

$$\min\left(X_i^{*(n)}, X_i\right) \leq Y_i^{(n)} \leq \max\left(X_i^{*(n)}, X_i\right).$$

Nous pouvons donc réécrire notre estimateur sous la forme

$$\widehat{\theta}_n^{M1} = \frac{1}{n}\sum_{i=1}^{n}\left(g(X_i) - g'\left(Y_i^{(n)}\right)\left(\bar{X}_n - \mu_X\right)\right)$$

$$= \hat{\theta}_n - \left(\bar{X}_n - \mu_X\right)\frac{1}{n}\sum_{i=1}^{n} g'\left(Y_i^{(n)}\right)$$

$$\widehat{\theta}_n^{M1} = \underbrace{\hat{\theta}_n}_{\to \theta p.s.} - \underbrace{\left(\bar{X}_n - \mu_X\right)}_{\to 0 p.s.} \underbrace{\frac{1}{n}\sum_{i=1}^{n} g'\left(Y_i^{(n)}\right)}_{<\infty} \to \theta p.s.$$

puisque

1. La loi forte des grands nombres implique que

$$X_i^{*(n)} - X_i = \mu_X - \bar{X}_n \overset{n\to\infty}{\to} 0 p.s.$$

2. $Y_i^{(n)} - X_i \overset{n\to\infty}{\to} 0 p.s.$ puisque

$$\underbrace{\min\left(X_i^{*(n)} - X_i, 0\right)}_{\overset{n\to\infty}{\to} 0 p.s.} \leq Y_i^{(n)} - X_i \leq \underbrace{\max\left(X_i^{*(n)} - X_i, 0\right)}_{\overset{n\to\infty}{\to} 0 p.s.}.$$

3. Si g est une fonction continue, alors $g\left(Y_i^{(n)}\right) - g(X_i) \overset{n\to\infty}{\to} 0 p.s.$

4. Sous certaines conditions, nous aurons aussi

L'estimateur est biaisé mais asymptotiquement sans biais. En effet,

$$\begin{aligned}
&E\left[\widehat{\theta}_n^{M1}\right] \\
&= E\left[\hat{\theta}_n\right] - E\left[\left(\bar{X}_n - \mu_X\right)\frac{1}{n}\sum_{i=1}^{n} g'\left(Y_i^{(n)}\right)\right] \\
&= \theta - \underbrace{E\left[\left(\bar{X}_n - \mu_X\right) g'\left(Y_1^{(n)}\right)\right]}_{=biais}
\end{aligned}$$

$$
\begin{aligned}
|biais| &\leq E\left[\left|\left(\bar{X}_n - \mu_X\right) g'\left(Y_1^{(n)}\right)\right|\right] \\
&\leq \sqrt{E\left[\left(\bar{X}_n - \mu_X\right)^2\right]}\sqrt{E\left[\left(g'\left(Y_1^{(n)}\right)\right)^2\right]} \\
&\text{(inégalité de Cauchy-Schwartz)} \\
&\cong \sqrt{Var\left[\bar{X}_n\right]}\sqrt{E\left[(g'(X_1))^2\right]} \\
&= \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

Utile pour déterminer l'intervalle de confiance construit autour de l'estimateur.

$$\sqrt{n}\left(\widehat{\theta}_n^{M1} - \theta\right) = \underbrace{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}_{\overset{n\to\infty}{\Rightarrow} N(0,\sigma^2)} - \underbrace{\sqrt{n}\left(\bar{X}_n - \mu_X\right)}_{\overset{n\to\infty}{\Rightarrow} N(0,Var[X])} \underbrace{\frac{1}{n}\sum_{i=1}^{n} g'\left(Y_i^{(n)}\right)}_{\overset{n\to\infty}{\to} E[g'(X)] \text{ probabilité}}$$

puisque

1. $\sqrt{n}\left(\mu_X - \bar{X}_n\right) \overset{n\to\infty}{\Rightarrow} N(0, Var[X])$ (convergence en loi).
2. Si les deux premiers moments de $g'(X)$ existent, la loi forte des grands nombres implique que $\frac{1}{n}\sum_{i=1}^{n} g'(X_i)$ converge presque surement vers son espérance.
3. Sous certaines conditions techniques, nous aurons $E\left[\left|g'\left(Y_1^{(n)}\right) g'(X_1)\right|\right] \overset{n\to\infty}{\to}$, ce qui entraine que $\frac{1}{n}\sum_{i=1}^{n} g'\left(Y_i^{(n)}\right)$ converge en probabilité vers $E\left[g'(X)\right]$.

Rappel :

$$\sqrt{n}\left(\widehat{\theta}_n^{M1}-\theta\right)=\underbrace{\sqrt{n}\left(\hat{\theta}_n-\theta\right)}_{\overset{n\to\infty}{\Rightarrow}N(0,\sigma^2)}-\underbrace{\sqrt{n}\left(\bar{X}_n-\mu_X\right)}_{\overset{n\to\infty}{\Rightarrow}N(0,Var[X])}\underbrace{\frac{1}{n}\sum_{i=1}^{n}g'\left(Y_i^{(n)}\right)}_{\overset{n\to\infty}{\to}E[g'(X)]\text{ probabilité}}$$

- On remarque que la variance de l'estimateur dépendra aussi du deuxième terme qui occasionne dans bien des cas une réduction de la variance.

- Il est difficile d'estimer la variance de l'estimateur à l'aide d'un seul échantillon.
- Lavariance de l'estimateur est souvent estimée en utilisant la resimulation, c'est-à-dire que
  1. Un échantillon de taille n est simulé, produisant une estimation $\widehat{\theta}_{n,j}^{M1}$
  2. L'étape 1 est répétée m fois, générant un échantillon d'estimation $\widehat{\theta}_{n,1}^{M1}, \ldots, \widehat{\theta}_{n,m}^{M1}$.
  3. La variance de l'estimateur est estimée par la variance échantillonale de $\widehat{\theta}_{n,1}^{M1}, \ldots, \widehat{\theta}_{n,m}^{M1}$.
- Cette façon de procéder détruit le gain obtenu par la réduction de variance car le temps requis pour estimer la variance est de loin supérieur au temps requis pour effectuer une simulation de Monte Carlo naîve.

Appariement du premier moment II

Inconvénients

- Il est possible, dans certains cas, d'estimer la variance avec un seul échantillon (sans resimulation). Toutefois, le travail théorique requis est souvent lourd.

- Dans les modèles de marché n'admettant pas d'opportunité d'arbitrage, la valeur actualisée des actifs négociables sont des martingales sous une mesure neutre au risque.
- C'est cette idée qui est exploitée afin d'obtenir le premier moment théorique du processus simulé.
- De plus, la correction pour le premier moment est multiplicative au lieu d'être additive.

- Soit $\{X(t) : t \in \mathcal{T}\}$ le processus stochastique modélisant l'évolution de la valeur actualisée d'un actif négociable sous une mesure neutre au risque où T représente une famille d'indices temporels.
- Puisque $\{X(t) : t \in \mathcal{T}\}$ est une martingale, nous avons que

$$E[X(t)] = E[X(0)], \text{ pour tout } t \in \mathcal{T}.$$

- Lors de la simulation de la variable aléatoire $X(t)$, un échantillon $X_1(t), .., X_n(t)$ est crée. Or la moyenne échantillonnale n'est pas égale à la moyenne théorique, c'est-à-dire que

$$\bar{X}_n(t) = \frac{1}{n} \sum_{i=1}^{n} X_i(t) \neq E[X(0)].$$

- Duan et Simonato (1998) proposent la correction suivante :

$$\widehat{X}_i^{(n)}(t) \equiv X_i(t)\frac{E[X(0)]}{\bar{X}_n(t)}.$$

- On peut alors vérifier que

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{X}_i^{(n)}(t) = \frac{1}{n}\sum_{i=1}^{n}X_i(t)\frac{E[X(0)]}{\bar{X}_n(t)} = E[X(0)].$$

- L'estimateur est

$$\theta_n^{ME} \equiv \frac{1}{n}\sum_{i=1}^{n} g\left(\widehat{X}_i^{(n)}\right).$$

- Il est possible de montrer que :
  1. θ_n^{ME} converge presque surement vers θ lorsque $n \to \infty$;
  2. θ_n^{ME} est biaisé et le biais est de l'ordre de n^{-1} ;
  3. $\sqrt{n}\left(\theta_n^{ME} - \theta\right) \overset{n\to\infty}{\Rightarrow} N(0, \xi^2)$;
     1. Pour les options d'achat,

$$\xi^2 = Var^Q\left[e^{-rT} max(S_t - K, 0)\right] + \blacksquare^2 Var^Q\left[e^{-rT} S_t\right]$$
$$- 2\blacksquare Cov^Q\left[e^{-rT} max(S_t - K, 0), e^{-rT} S_t\right]$$
$$\text{où } \blacksquare = E^Q[S_t \mathbf{1}_{X>K}]/E\left[e^{rT} X(0)\right].$$

  4. un estimateur de la variance est donné, il est obtenu à l'aide du même échantillon que celui permettant d'estimer θ (pas de ré-échantillonnage) ;
  5. tous les cas étudiés démontrent une bonne réduction de la variance.

- On peut refaire l'exercice en utilisant

$$X_i^{**(n)} \equiv \left(X_i - \bar{X}_n\right) \frac{\sqrt{Var[X]}}{S_n} + \mu_X.$$

où $S_n^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (Xi - \bar{X}_n)^2$.

- Ainsi, les deux premiers moments empiriques de l'échantillon $X_1^{**(n)}, ...X_n^{**(n)}$ seront égaux à leur contrepartie théorique :

$$\overline{X}_n^{**(n)} = \frac{1}{n}\sum_{i=1}^{n} X_i^{**(n)} = \left(\bar{X}_n - \bar{X}_n\right)\frac{\sqrt{Var[X]}}{S_X^{(n)}} + \mu_X = \mu_X$$

$$\begin{aligned}(S_n^{**(n)})^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i^{**(n)} - \overline{X}_n^{**(n)}\right)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left((X_i - \bar{X}_n)\frac{\sqrt{Var[X]}}{S^{(n)}}\right)^2 \\ &= \frac{\sqrt{Var[X]}}{S_n^2}\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \\ &= Var[X]\end{aligned}$$

- $$\widehat{\theta}_n^{ME2} = \frac{1}{n}\sum_{i=1}^{n} g\left(X_i^{**(n)}\right)$$
- Le problème de la dépendance entre les trajectoires demeure toujours présent et il faut travailler un peu plus fort pour obtenir les propriétés asymptotiques de l'estimateur
- Barraquand (1995) a proposé une méthode dans laquelle il apparie les moments des variables aléatoires gaussiennes simulées dans bien des cas en finance.

Références

- J.Barraquand, 1995, "Numerical valuation of high dimensional multivariate european securities, Management Science, 41, 1882-1891.
- Phelim P. Boyle, Marc Broadie and Paul Glasserman, 1997, "Monte Carlo methods for security pricing", ournal of Economic dynamics and control,21, 1267-1321.
- Jin-Chuan Duan et Jean-Guy Simonato, 1998, "Empirical martingale simulation for asset prices", Management Science, 44, 1218-1233.
- Jin-Chuan Duan, Geneviève Gauthier et Jean-Guy Simonato, 2001, "Asymptotic Distribution of the EMS Option Price Estimator",Management Science

Travail pratique 3

Reprendre la simulation de l'exercice précédent (dans le contexte NGARCH) en utilisant l'appariement des moments.

1. Déterminer le prix des options d'achat d'une maturité d'un an et dont les prix d'exercice sont respectivement $0.9 \times S_0$, S_0 et $1.1 \times S_0$ à l'aide d'une simulation de Monte Carlo pour des tailles d'échantillon égales à 5000, 10000, 50 000, 100 000 et 500 000.
2. Construire les intervalles de confiance autour de ces prix.
3. Noter les temps de calcul nécessaires pour chaque taille d'échantillon.
4. Produisez un graphe présentant la marge d'erreur en abscisse et le temps de calcul en ordonnée.
5. Comparer vos résultats avec ceux obtenus sans réduction de la variance et avec l'utilisation de la variable antithétique et variable de contrôle (présenter les courbes sur le même graphe) et commentez.