Estimation de la valeur-à-risque

6-601-09 Simulation Monte Carlo

Geneviève Gauthier

HEC Montréal

Référence

Ces notes de cours sont fortement inspirées de

1. P. Glasserman : *Monte Carlo Methods in Financial Engineering* (2004), chapitre 9.

Notation I

Introduction

1. Δ_t est la longueur de la période de temps considérée. Nous l'appelerons l'*horizon*.
2. $\mathbf{S}_t$ est un vecteur contenant les m facteurs de risque (prix, taux d'intérêt, taux de change, volatilités, etc.) au temps t
   2.1 $\Delta\mathbf{S} = \mathbf{S}_{t+\Delta_t} - \mathbf{S}_t$ est la variation des facteurs de risque au cours de l'horizon Δ_t.
3. $V(\mathbf{S}, t)$ est la valeur du portefeuille au temps t étant donné que $\mathbf{S} = \mathbf{S}_t$.
   3.1 $\Delta V = V(\mathbf{S}_{t+\Delta_t}, t + \Delta_t) - V(\mathbf{S}_t, t)$ est la variation de la valeur du portefeuille au cours de l'horizon Δ_t.
4. $L = -\Delta V$ est la perte au cours de l'horizon Δ_t. (positif pour une perte, négatif pour un gain)

Notation II

Introduction

5. $F_L(x) = P(L \leq x)$ est la fonction de répartition de la perte.
6. En supposant que L ait une distribution continue, la valeur-à-risque (VaR) de niveau α, $0 < \alpha < 1$, est la quantité x_α telle que

$$1 - F_L(x_\alpha) = P(L > x_\alpha) = \alpha$$

7. Dans ce qui suit, nous supposerons que l'instant présent est représenté par $t = 0$.

Monte Carlo Naif

Algorithme I

Monte Carlo naïf

1. Simuler n scénarios indépendants et identiquement distribués $\Delta\mathbf{S}^{(1)}, \ldots, \Delta\mathbf{S}^{(n)}$ pour les facteurs de risque au temps Δ_t.
2. Créer les n valeurs de portefeuille : $V^{(1)}, \ldots, V^{(n)}$ où $V^{(i)} = V\left(\mathbf{S}^{(i)}, \Delta_t\right)$
3. Créer les scénarios de perte : $L^{(1)}, \ldots, L^{(n)}$.

Algorithme II
Monte Carlo naïf

4. Pour diverses valeurs de x, il est possible d'estimer la probabilité $p_x = P[L \geq x]$ par la proportion échantillonnale

$$\widehat{p}_x = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{L^{(i)} > x}.$$

   4.1 $z_{\frac{\beta}{2}}$ est le nombre tel que $P\left(Z > z_{\frac{\beta}{2}}\right) = \frac{\beta}{2}$ où Z est $N(0,1)$.

Algorithme III

Monte Carlo naïf

5. Parce que les scénarios sont iid, le théorème central limite s'applique et il est possible de calculer un intervalle de confiance de niveau $1-\beta$ pour p_x :

$$\left[\widehat{p}_x - z_{\frac{\beta}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n}};\ \widehat{p}_x + z_{\frac{\beta}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n}}\right]$$

   5.1 Cet intervalle est basé sur l'approximation de la loi binomiale par la loi normale. Il faut donc que n soit assez grand, d'autant plus que p_x est proche de zéro ou de un.

   5.2 Cette façon de procéder diffère de l'estimation d'un quantile présentée au début du trimestre. Ici, il ne s'agit pas d'estimer x mais d'estimer p_x.

Algorithme IV

Monte Carlo naïf

6. La VaR de niveau α s'estime par inversion de la fonction de répartition empirique de L. Elle peut aussi s'estimer par interpolation (Glasserman, p.490).
   6.1 Pour l'interpolation linéaire, supposons que $\widehat{p}_{x_1} < \alpha < \widehat{p}_{x_2}$., Alors $\widehat{x}_\alpha = x_1 + \frac{\alpha - \widehat{p}_{x_1}}{\widehat{p}_{x_2} - \widehat{p}_{x_1}} (x_2 - x_1)$.
   6.2 Si on a plusieurs couples $(x_1, \widehat{p}_{x_1})$, ..., $(x_k, \widehat{p}_{x_k})$, il est possible d'utiliser d'autres méthodes d'interpolation basées sur des polynomes, sur des splines, etc.
7. Il est possible de calculer une marge d'erreur (Glasserman, p.490)...En supposant que L admette une fonction de densité f_L positive dans le voisinage de x_p, alors

$$\sqrt{n}\left(\widehat{x}_p - x_p\right) \overset{n\to\infty}{\Rightarrow} \frac{\sqrt{p(1-p)}}{f_L(x_p)} N(0,1).$$

Algorithme V
Monte Carlo naïf

7.1 **Idée intuitive** : Supposons que nous voulions estimer $G(\theta)$ où G est une fonction suffisament lisse et que nous possédions un estimateur $\widehat{\theta}$ de θ telle que $\sqrt{n}(\widehat{x}_p - x_p) \overset{n\to\infty}{\Rightarrow} \sigma N(0,1)$. Alors, le développement en série de Taylor de $G(x)$ autour de $G(\theta)$ donne

$$\sqrt{n}\left(G\left(\widehat{\theta}\right) - G(\theta)\right) = G'(\theta) \underbrace{\sqrt{n}\left(\widehat{\theta} - \theta\right)}_{\overset{n\to\infty}{\Rightarrow}\sigma N(0,1)} + \frac{1}{2}\underbrace{G''(\xi)}_{\to G''(\theta)} \underbrace{\sqrt{n}\left(\widehat{\theta} - \theta\right)^2}_{\to 0}$$

où ξ est compris entre $\widehat{\theta}$ et θ. Dans le cas de la VaR, $G = F_L^{-1}$ d'où $G' = \frac{1}{f_L}$.

Goulot d'étranglement

Monte Carlo naif

1. Une des principales difficultés rencontrées avec ce dernier algorithme est que l'étape 2 (*évaluation de la valeur du portefeuille*) peut être très longue lorsque les portefeuilles contiennent une grande quantité de titres et/ou de produits dérivés.
   1.1 Parfois, il faut évaluer les produits dérivés à l'aide d'une simulation de Monte Carlo (en monde neutre au risque) imbriquée dans la simulation principale.
2. Dans ces cas, le nombre de trajectoires doit être limité, ce qui implique une marge d'erreur relativement grande (d'autant plus que nous devons estimer un événement rare).
3. Une façon de pallier à cette difficulté est d'introduire des méthodes de réduction de la variance.

Approximation quasi-analytique

Approximation Delta-Gamma

1. Dans ce qui suit, nous allons présenter une approximation de la distribution de la perte L à l'aide de son développement en série de Taylor.
2. Dans certains cas, lorsque la valeur du portefeuille est bien approchée par la tendance et la convexité, cette approximation sera suffisante.
3. Par contre, dans la majorité des cas, l'approximation ne sera pas assez précise pour être utilisée seule mais elle pourra être utilisée comme variable de contrôle afin de réduire substantiellement la marge d'erreur en petit échantillon.

Approximation Delta-Gamma

Fondements

Approximation Delta-Gamma

- Rappelons que $\Delta V = V\left(\mathbf{S}_{\Delta_t}, \Delta_t\right) - V\left(\mathbf{S}_0, 0\right).$
- En utilisant le développement en série de Taylor de ΔV autour de $\left(\mathbf{S}_0, 0\right)$, nous obtenons

$$
\begin{aligned}
\Delta V &= V\left(\mathbf{S}_{\Delta_t}, \Delta_t\right) - V\left(\mathbf{S}_0, 0\right) \\
&\cong \frac{\partial V}{\partial \Delta_t}\left(\mathbf{S}_0, 0\right)\Delta_t + \sum_{j=1}^{m}\left(\frac{\partial V}{\partial S_j}\left(\mathbf{S}_0, 0\right)\right)\left(S_{j,\Delta_t} - S_{j,0}\right) \\
&\quad + \frac{1}{2}\sum_{j=1}^{m}\sum_{j^*=1}^{m}\left(\frac{\partial^2 V}{\partial S_j \partial S_{j^*}}\left(\mathbf{S}_0, 0\right)\right)\left(S_{j,\Delta_t} - S_{j,0}\right)\left(S_{j^*,\Delta_t} - S_{j^*,0}\right) \\
&= \frac{\partial V}{\partial \Delta_t}\Delta_t + \delta^\top \Delta\mathbf{S} + \frac{1}{2}\Delta\mathbf{S}^\top \Gamma \Delta\mathbf{S}
\end{aligned}
$$

où $\delta = \left(\frac{\partial V}{\partial S_1}, \ldots \frac{\partial V}{\partial S_m}\right)^\top$ et $\Gamma_{ij} = \frac{\partial^2 V}{\partial S_j \partial S_{j^*}}$ et toutes les dérivées sont évaluées au point initial $\left(\mathbf{S}_0, 0\right).$

Diagonalisation I

Approximation Delta-Gamma

1. Dans le cas où la variation des facteurs de risque est de loi normale multivariée, disons $\Delta \mathbf{S} = N(\mathbf{0}, \Sigma_{\mathbf{S}})$ alors

$$\Delta \mathbf{S} = \mathbf{C}\mathbf{Z}$$

où

1.1 $\mathbf{Z}$ est $N(\mathbf{0}, \mathbf{I}_m)$,
1.2 $\mathbf{I}_m$ est la matrice identité de dimension m et
1.3 $\mathbf{C}\mathbf{C}^\top = \Sigma_{\mathbf{S}}$.

Diagonalisation II

Approximation Delta-Gamma

2. La perte peut être approximée par

$$
\begin{aligned}
L &= -\Delta V \\
&\cong -\frac{\partial V}{\partial \Delta_t}\Delta_t - \delta^\top \Delta \mathbf{S} - \frac{1}{2}\Delta \mathbf{S}^\top \Gamma \Delta \mathbf{S} \\
&= -\frac{\partial V}{\partial \Delta_t}\Delta_t - \delta^\top \mathbf{CZ} - \frac{1}{2}(\mathbf{CZ})^\top \Gamma (\mathbf{CZ}) \\
&= a - \left(\mathbf{C}^\top \delta\right)^\top \mathbf{Z} - \frac{1}{2}\mathbf{Z}^\top \mathbf{C}^\top \Gamma \mathbf{C} \mathbf{Z}
\end{aligned}
$$

où $a = -\frac{\partial V}{\partial \Delta_t}\Delta_t$ est une constante.

Diagonalisation III
Approximation Delta-Gamma

3. Il devient intéressant de diagonaliser le terme quadratique de l'équation de perte.
   3.1 Soit $\widetilde{\mathbf{C}}$, une matrice carrée satisfaisant $\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^\top = \Sigma_{\mathbf{S}}$ (décomposition de Choleski, par exemple)
   3.2 La matrice $-\frac{1}{2}\widetilde{\mathbf{C}}^\top \Gamma \widetilde{\mathbf{C}}^\top$ est symétrique et admet donc la représentation $-\frac{1}{2}\widetilde{\mathbf{C}}^\top \Gamma \widetilde{\mathbf{C}}^\top = \mathbf{U}\Lambda\mathbf{U}^\top$ où
      3.2.1 Λ est une matrice diagonale formée des valeurs propres $\lambda_1, ..., \lambda_m$ de $\Sigma_{\mathbf{S}}$
      3.2.2 et $\mathbf{U}$ est une matrice dont les colonnes sont composées des vecteurs propres de $-\frac{1}{2}\widetilde{\mathbf{C}}^\top \Gamma \widetilde{\mathbf{C}}^\top$ et satisfaisant $\mathbf{U}\mathbf{U}^\top = \mathbf{I}_m$.
   3.3 Posons $\mathbf{C} = \widetilde{\mathbf{C}}\mathbf{U}$ et notons que
      3.3.1 $\mathbf{C}\mathbf{C}^\top = \widetilde{\mathbf{C}}\mathbf{U}\mathbf{U}^\top\widetilde{\mathbf{C}}^\top = \widetilde{\mathbf{C}}\mathbf{I}_m\widetilde{\mathbf{C}}^\top = \Sigma_{\mathbf{S}}$.
      3.3.2 $-\frac{1}{2}\mathbf{C}^\top\Gamma\mathbf{C} = -\frac{1}{2}\mathbf{U}^\top\left(\widetilde{\mathbf{C}}^\top\Gamma\widetilde{\mathbf{C}}\right)\mathbf{U} = \mathbf{U}^\top\left(\mathbf{U}\Lambda\mathbf{U}^\top\right)\mathbf{U} = \Lambda$.

Diagonalisation IV

Approximation Delta-Gamma

3.4 En posant $\mathbf{b} = -\mathbf{C}^\top \delta = -\mathbf{C}^\top \left(\frac{\partial V}{\partial S_1}, \ldots \frac{\partial V}{\partial S_m} \right)^\top$, l'équation de perte devient

$$\begin{aligned} L &\cong a - \left(\mathbf{C}^\top \delta \right)^\top \mathbf{Z} - \frac{1}{2} \mathbf{Z}^\top \mathbf{C}^\top \Gamma \mathbf{C} \mathbf{Z} \\ &= a + \mathbf{b}^\top \mathbf{Z} + \mathbf{Z}^\top \Lambda \mathbf{Z} \\ &= a + \sum_{i=1}^{m} \left(b_i Z_i + \lambda_i Z_i^2 \right) \\ &= Q \end{aligned}$$

3.5 $P[L > x] \cong P[Q > x]$.

3.5.1 Nous allons déterminer la distribution de Q afin de calculer analytiquement le membre de droite de l'approximation ci-dessus.

Exemple I

Approximation Delta-Gamma

1. Considérons un portefeuille contenant 100 parts de 4 options différentes portant sur deux titres sous-jacents dont l'évolution des prix, en monde neutre au risque, sont données par
   1.1 Titre A : $dS_t^A = rS_t^A dt + \sigma_A S_t^A dW_t^A$.
   1.2 Titre B : $dS_t^B = rS_t^B dt + \sigma_B S_t^B dW_t^B$.
2. Supposons que $r = 5\%$, $S_0^A = 50$, $\sigma_A = 20\%$, $S_0^B = 75$, $\sigma_B = 30\%$ et $\rho = \mathrm{Corr}\left[W_t^A, W_t^B\right] = 50\%$.
3. Les options sont :
   3.1 Option d'achat du titre A : $K_1 = 45$, $T_1 = 0.25$.
   3.2 Option de vente du titre A : $K_2 = 50$, $T_2 = 1$.
   3.3 Option d'achat du titre B : $K_3 = 80$, $T_3 = 0.5$.
   3.4 Option de vente du titre B : $K_4 = 75$, $T_4 = 0.25$.
4. L'horizon de calcul de la VaR est une semaine : $\Delta_t = \frac{1}{50} = 0.02$.

Exemple II

Approximation Delta-Gamma

5. La valeur du portefeuille au temps $t < \min\left(T_1, ..., T_4\right)$ est

$$\begin{aligned} V\left(\mathbf{S}_t, t\right) &= w_1 C\left(S_t^A, t, K_1, T_1\right) + w_2 P\left(S_t^A, t, K_2, T_2\right) \\ &\quad + w_3 C\left(S_t^B, t, K_3, T_3\right) + w_4 P\left(S_B, t, K_4, T_4\right) \end{aligned}$$

où $w_1, ..., w_4$ sont les nombres d'options détenues où

5.1 $C\left(S_t, t, K, T\right) =$ valeur au temps t d'une option d'achat d'échéance t et de prix d'exercice K,

5.2 $P\left(S_t, t, K, T\right) =$ valeur au temps t d'une option de vente d'échéance t et de prix d'exercice K.

Exemple III

Approximation Delta-Gamma

6. Nous avons deux facteurs de risques.
   6.1 Parce que les prix sont de loi lognormale, nous devons les transformer afin d'obtenir des facteurs de risque gaussiens d'espérance nulle. :
   6.2 $\mathbf{X}_{\Delta_t} = \begin{pmatrix} \ln S^A_{\Delta_t} - \left(r - \frac{1}{2}\sigma_A^2\right)\Delta_t \\ \ln S^B_{\Delta_t} - \left(r - \frac{1}{2}\sigma_B^2\right)\Delta_t \end{pmatrix}$
   6.3 $\Delta\mathbf{X} = \begin{pmatrix} \ln S^A_{\Delta_t} - \left(r - \frac{1}{2}\sigma_A^2\right)\Delta_t - \ln S^A_0 \\ \ln S^B_{\Delta_t} - \left(r - \frac{1}{2}\sigma_B^2\right)\Delta_t - \ln S^B_0 \end{pmatrix} \sim N\left(\mathbf{0}, \Sigma_{\Delta X}\right)$
   6.4 où $\Sigma_{\Delta X} = \begin{pmatrix} \sigma_A^2 & \sigma_A\sigma_B\rho \\ \sigma_A\sigma_B\rho & \sigma_B^2 \end{pmatrix} = \begin{pmatrix} 0.04 & 0.03 \\ 0.03 & 0.09 \end{pmatrix}$.

Exemple IV

Approximation Delta-Gamma

7. De façon plus générale, si nous avons l'EDS

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

et que la solution de cette équation n'est pas gaussienne, nous devons la transformer.
Si f est une fonction deux fois continûment différentiable, alors le lemme d'Itô implique que

$$\begin{aligned} & df(t, S_t) \\ = \;& \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial S^2} d\langle S\rangle_t \\ = \;& \left(\frac{\partial f}{\partial t} + \mu(t, S_t)\frac{\partial f}{\partial S} + \frac{\sigma^2(t, S_t)}{2}\frac{\partial^2 f}{\partial S^2}\right) dt + \sigma(t, S_t)\frac{\partial f}{\partial S} dW_t. \end{aligned}$$

Pour que $f(S_t, t)$ soit gaussienne d'espérance nulle, il faut que $\sigma(t, S_t)\frac{\partial f}{\partial S} = \sigma_t$ soit une fonction déterministe du temps et que $\frac{\partial f}{\partial t} + \mu(t, S_t)\frac{\partial f}{\partial S} + \frac{\sigma^2(t,S_t)}{2}\frac{\partial^2 f}{\partial S^2} = 0$

Exemple V
Approximation Delta-Gamma

Selon le système d'EDS à résoudre, il est possible de solutionner le système d'équations différentielles non-stochastiques

$$\begin{aligned} \frac{\partial f}{\partial S}(t,S) &= \frac{\sigma_t}{\sigma(t,S)}, \\ \frac{\partial f}{\partial t}(t,S) - \frac{\sigma_t}{2}\frac{\partial \sigma}{\partial S}(t,S) &= -\mu(t,S)\frac{\sigma_t}{\sigma(t,S)} \end{aligned}$$

afin de déterminer la transformation à appliquer aux variables.

Exemple VI

Approximation Delta-Gamma

8. Calcul de δ

   8.1 $\frac{\partial C}{\partial S}(S_t, t) = N(d_t)$
   8.2 $\frac{\partial P}{\partial S}(S_t, t) = N(d_t) - 1.$

      8.2.1 $d_t = \left(\ln S_t - \ln K + \left(r + \frac{\sigma_A^2}{2}\right)(T - t)\right) / \sigma\sqrt{(T-t)}$
      8.2.2 K est le prix d'exercice,
      8.2.3 T est l'échéance,
      8.2.4 $N(\bullet)$ est la fonction de répartition d'une $N(0,1)$.

   8.3 $\delta = \left(\frac{\partial V}{\partial X^A}; \frac{\partial V}{\partial X^B}\right)^\top = (2636.1029; 231.7058)^\top$

      8.3.1 $\frac{\partial V}{\partial X} = \frac{\partial V}{\partial S}\frac{\partial S}{\partial X} = S\frac{\partial V}{\partial S}$ (règle de dérivées en chaîne)
      8.3.2 $\frac{\partial S}{\partial X} = \frac{\partial}{\partial X} e^{X+c} = e^{X+c} = S.$
      8.3.3 $\frac{\partial V}{\partial X^A} = S^A \left[w_1 N(d_1) + w_2 (N(d_2) - 1)\right]$
      8.3.4 $\frac{\partial V}{\partial X^B} = S^B \left[w_3 N(d_3) + w_4 (N(d_4) - 1)\right]$

Exemple VII
Approximation Delta-Gamma

9. Calcul de $\Gamma = \begin{pmatrix} 16865.43 & 0 \\ 0 & 11243.73 \end{pmatrix}$

   9.1 $\frac{\partial^2 C}{\partial S^2} = \frac{\partial^2 P}{\partial S^2} = \frac{1}{\sigma\sqrt{T}} \frac{1}{S} n(d)$ où n est la fonction de densité d'une variable aléatoire $N(0,1)$.

   9.2 $\frac{\partial^2 C}{\partial X^2} = \frac{\partial}{\partial X}\left(\frac{\partial C}{\partial X}\right) = \frac{\partial}{\partial X}\left(S\frac{\partial C}{\partial S}\right) = \frac{\partial S}{\partial X}\frac{\partial}{\partial S}\left(S\frac{\partial C}{\partial S}\right) = \frac{\partial S}{\partial X}\left(\frac{\partial C}{\partial S} + S\frac{\partial^2 C}{\partial S^2}\right) = S\left(\frac{\partial C}{\partial S} + S\frac{\partial^2 C}{\partial S^2}\right) = S\left(N(d) + \frac{n(d)}{\sigma\sqrt{T}}\right)$

   9.3 $\frac{\partial^2 P}{\partial X^2} = S\left(N(d) - 1 + \frac{n(d)}{\sigma\sqrt{T}}\right)$

Exemple VIII

Approximation Delta-Gamma

10. $\widetilde{\mathbf{C}} = \begin{pmatrix} 0.2 & 0 \\ 0.15 & 0.259\,81 \end{pmatrix}$ (Décomposition de Choleski de $\Sigma_{\Delta X}$).
11. Diagonalisation $(-\frac{1}{2}\widetilde{\mathbf{C}}^\top \Gamma \widetilde{\mathbf{C}} = \mathbf{U}\Lambda\mathbf{U}^\top)$:

    11.1 $-\frac{1}{2}\widetilde{\mathbf{C}}^\top \Gamma \widetilde{\mathbf{C}} = \begin{pmatrix} -472.80 & -218.08 \\ -218.08 & -377.74 \end{pmatrix}$

    11.2 $\Lambda = \begin{pmatrix} -202.07 & 0 \\ 0 & -648.47 \end{pmatrix}$

    11.3 $\mathbf{U} = \begin{pmatrix} 0.627\,32 & 0.778\,76 \\ -0.778\,76 & 0.627\,32 \end{pmatrix}$
12. $\mathbf{C} = \widetilde{\mathbf{C}}\mathbf{U} = \begin{pmatrix} 0.125\,46 & 0.155\,75 \\ -0.108\,23 & 0.279\,80 \end{pmatrix}$
13. $\mathbf{b} = -\mathbf{C}^\top \delta = \begin{pmatrix} 314.71 \\ 465.4 \end{pmatrix}$
14. $Q = a + \sum_{i=1}^{n} \left(b_i Z_i + \lambda_i Z_i^2\right)$

Distribution de Q I

Approximation Delta-Gamma

1. La distribution de Q est liée à sa fonction génératrice des moments :

$$\begin{aligned} \mathrm{E}\left[e^{\theta Q}\right] &= \mathrm{E}\left[e^{\theta\left(a+\sum_{i=1}^{m}\left(b_i Z_i+\lambda_i Z_i^2\right)\right)}\right] \\ &= e^{a\theta}\prod_{i=1}^{m}\mathrm{E}\left[e^{\left(b_i Z_i+\lambda_i Z_i^2\right)\theta}\right] \\ &= e^{a\theta}\prod_{i=1}^{m}e^{\psi_i(\theta)} \end{aligned}$$

où $\psi_i(\theta) = \ln e^{\psi_i(\theta)} = \ln \mathrm{E}\left[e^{\left(b_i Z_i+\lambda_i Z_i^2\right)\theta}\right]$.

2. Si $\lambda_i = 0$, alors

$$\psi_i(\theta) = \ln e^{\psi_i(\theta)} = \ln \mathrm{E}\left[e^{b_i\theta Z_i}\right] = \ln e^{\frac{1}{2}\theta^2 b_i^2} = \frac{1}{2}\theta^2 b_i^2.$$

Distribution de Q II

Approximation Delta-Gamma

3. Si $\lambda_i \neq 0$, alors $b_i Z_i + \lambda_i Z_i^2 = \lambda_i \left(Z_i + \frac{b_i}{2\lambda_i} \right)^2 - \frac{b_i^2}{4\lambda_i}$ est une transformation linéaire d'une variable de loi du khi-deux non-centrale.

   3.1 Il est possible de montrer (Glasserman, p.487) que si $\theta < \frac{1}{2}$, alors

   $$\mathrm{E}\left[e^{\theta (Z_i - c)^2} \right] = \frac{1}{\sqrt{1 - 2\theta}} \exp\left(\frac{\theta c^2}{1 - 2\theta} \right).$$

   3.2 La contrainte sur θ n'est pas problématique puisque la fonction génératrice des moments est généralement utilisée lorsque θ est dans le voisinage de zéro.

Distribution de Q III

Approximation Delta-Gamma

3.3 Par conséquent, si $\lambda_i \theta < \frac{1}{2}$, alors

$$
\begin{aligned}
\psi_i(\theta) &= \ln e^{\psi_i(\theta)} \\
&= \ln \mathrm{E}\left[e^{\left(b_i Z_i + \lambda_i Z_i^2\right)\theta}\right] \\
&= \ln e^{-\frac{b_i^2}{4\lambda_i}\theta} \mathrm{E}\left[e^{\lambda_i \left(Z_i + \frac{b_i}{2\lambda_i}\right)^2 \theta}\right] \\
&= \ln e^{-\frac{b_i^2}{4\lambda_i}\theta} \frac{1}{\sqrt{1-2\lambda_i\theta}} \exp\left(\frac{\lambda_i\theta\left(\frac{b_i}{2\lambda_i}\right)^2}{1-2\lambda_i\theta}\right) \\
&= \frac{1}{2}\left[\frac{\theta^2 b_i^2}{1-2\theta\lambda_i} - \ln\left(1-2\lambda_i\theta\right)\right]
\end{aligned}
$$

Distribution de Q IV

Approximation Delta-Gamma

4. $\mathrm{E}\left[e^{\theta Q}\right] = e^{a\theta} \prod_{i=1}^{m} \exp \psi_i(\theta) = \exp \psi(\theta)$ où

$$\begin{aligned} \psi(\theta) &= a\theta + \sum_{i=1}^{m} \psi_i(\theta) \\ &= a\theta + \frac{1}{2} \sum_{i=1}^{m} \left[\frac{\theta^2 b_i^2}{1 - 2\theta\lambda_i} - \ln(1 - 2\lambda_i \theta) \right]. \end{aligned}$$

Distribution de Q V

Approximation Delta-Gamma

5. La fonction caractéristique est $\mathrm{E}\left[e^{iuQ}\right]$.

   5.1 Suivant un cheminement analogue à celui du calcul de la fonction génératrice des moments,

   $$\phi(u) = \mathrm{E}\left[e^{iuQ}\right] = \exp\psi(iu).$$

   5.2 Il existe un lien entre la fonction caractéristique d'une variable aléatoire et sa fonction de répartition (Glasserman, p.487) :

   $$\begin{aligned} & P[Q \leq x] - P[Q \leq x - y] \\ = & \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\phi(u)\left[\frac{e^{iu\,y} - 1}{iu}\right]e^{-iux}\right)du \end{aligned}$$

   5.3 L'obtention de $P[Q \leq x]$ s'obtient en choisissant y suffisamment grand pour que $P[Q \leq x - y]$ devienne négligeable.

   5.4 Les références pour la solution numérique se trouvent dans Glasserman, p.488.

Variable de contrôle

Variable de contrôle

Introduction

Approximation Delta-Gamma comme variable de contrôle

Lorsque l'approximation Delta-Gamma est trop imprécise pour être utilisée directement, il est possible de s'en servir comme variable de contrôle.

Algorithme I

Approximation Delta-Gamma comme variable de contrôle

1. Simuler n scénarios indépendants et identiquement distribués $\Delta\mathbf{S}^{(1)}, ..., \Delta\mathbf{S}^{(n)}$ pour les facteurs de risque au temps Δ_t en utilisant la décomposition $\mathbf{CZ}^{(j)}$
   1.1 Créer les n valeurs de portefeuille : $V^{(1)}, ..., V^{(n)}$ où $V^{(j)} = V\left(\mathbf{S}^{(j)}, \Delta_t\right)$.
   1.2 Pour chacune des trajectoires, calculer $Q^{(j)} = a + \sum_{i=1}^{n}\left(b_i Z_i^{(j)} + \lambda_i\left(Z_i^{(j)}\right)^2\right)$.
2. Créer les scénarios de perte : $L^{(1)}, ..., L^{(n)}$.
3. Pour diverses valeurs de x, il est possible d'estimer la probabilité $p_x = P\left[L \geq x\right]$ par la proportion échantillonnale

$$\widehat{\widehat{p}}_x = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{L^{(i)}>x} - \widehat{\beta}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{Q^{(i)}>x} - P\left(Q > x\right)\right)$$

Algorithme II

Approximation Delta-Gamma comme variable de contrôle

3.1 où $z_{\frac{\beta}{2}}$ est le nombre tel que $P\left(Z > z_{\frac{\beta}{2}}\right) = \frac{\beta}{2}$ où Z est $N(0,1)$

3.2 $\widehat{\beta}$ est l'estimation du coefficient β minimisant la variance (voir le document sur les variables de contrôle)

4. Parce que les scénarios sont iid, le théorème central limite s'applique et il est possible de calculer un intervalle de confiance de niveau $1-\beta$ pour p_x :

$$\left[\widehat{\widehat{p}}_x - z_{\frac{\beta}{2}}\sqrt{\frac{\widehat{\widehat{p}}_x\left(1-\widehat{\widehat{p}}_x\right)}{n}} ; \ \widehat{\widehat{p}}_x + z_{\frac{\beta}{2}}\sqrt{\frac{\widehat{\widehat{p}}_x\left(1-\widehat{\widehat{p}}_x\right)}{n}}\right]$$

4.1 Cet intervalle est basé sur l'approximation de la loi binomiale par la loi normale. Il faut donc que n soit assez grand, d'autant plus que p_x est proche de zéro ou de un.

Échantillonnage stratégique

Introduction I

Échantillonnage stratégique

- Rappelons qu'avec la variable de contrôle, nous avions

$$\widehat{\widehat{p}}_x = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{L^{(i)}>x} - \widehat{\beta}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{Q^{(i)}>x} - P\left(Q>x\right)\right)$$

- Le problème est que si x est grand, nous aurons beaucoup de trajectoires pour lesquelles $\mathbf{1}_{L^{(i)}>x} = 0$, réduisant de ce fait même la portée de la variable de contrôle.
- Utilisons donc l'échantillonnage stratégique. Cependant, la distribution de la perte est une fonction complexe des facteurs de risque rendant très difficile le choix de la nouvelle densité. L'idée maîtresse est de se servir de la connaissance de la distribution de Q pour bien choisir cette densité.

Introduction II
Échantillonnage stratégique

$L \cong Q = a + \sum_{i=1}^{m} b_i Z_i + \sum_{i=1}^{m} \lambda_i Z_i^2$ où les Z_i sont iid $N\left(0, \sigma^2\right)$.

Intuitivement, les pertes importantes surviendront lorsque

1. Z_i est grand et positif et $b_i > 0$,
2. Z_i est grand et négatif et $b_i < 0$,
3. Z_i^2 est grand et $\lambda_i > 0$.

La solution intuitive est

1. d'augmenter l'espérance de Z_i lorsque $b_i > 0$,
2. diminuer l'espérance de Z_i lorsque $b_i < 0$,
3. augmenter la variance de Z_i lorsque $\lambda_i > 0$.

Ce qui suit présente une façon de réaliser ce qui est proposé.

"Exponential twisting" I

Échantillonnage stratégique

1. Restreignons-nous à la famille de mesures de probabilité $\{P_\theta : \theta \in \Theta\}$ dont la dérivée de Radon-Nikodym est

$$\frac{dP_\theta}{dP} = \exp\left(\theta Q - \psi\left(\theta\right)\right)$$

2.

$$\begin{aligned} P\left[Q > x\right] &= \mathrm{E}\left[\mathbf{1}_{Q>x}\right] \\ &= \mathrm{E}_\theta\left[\frac{dP}{dP_\theta}\mathbf{1}_{Q>x}\right] \\ &= \mathrm{E}_\theta\left[e^{-\theta Q + \psi(\theta)}\mathbf{1}_{Q>x}\right] \end{aligned}$$

où $\mathrm{E}_\theta\left[\bullet\right]$ dénote l'espérance sous la nouvelle mesure P_θ.

"Exponential twisting" II

Échantillonnage stratégique

3. Un estimateur pour $P\left[L > x\right]$ est de simuler Q à l'aide de sa distribution sous P_θ et d'utiliser la moyenne échantillonnale

$$\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\theta Q^{(j)} + \psi\left(\theta\right)\right)\mathbf{1}_{L^{(j)}>x}$$

où $L^{(j)}$ et $Q^{(j)}$ sont issus de la j ième trajectoires.

3.1 $\frac{dP_\theta}{dP} = \exp\left(\theta Q - \psi\left(\theta\right)\right)$ agit comme facteur de correction pour avoir utiliser la "mauvaise" distribution.

"Exponential twisting" III

Échantillonnage stratégique

4. Le deuxième moment de $e^{-\theta Q+\psi(\theta)}\mathbf{1}_{L>x}$ est

$$
\begin{aligned}
& \mathrm{E}_\theta\left[\exp\left(-2\theta Q+2\psi\left(\theta\right)\right)\mathbf{1}_{L>x}\right] \\
= \; & \mathrm{E}\left[\frac{dP_\theta}{dP}\exp\left(-2\theta Q+2\psi\left(\theta\right)\right)\mathbf{1}_{L>x}\right] \\
= \; & \mathrm{E}\left[\exp\left(\theta Q-\psi\left(\theta\right)\right)\exp\left(-2\theta Q+2\psi\left(\theta\right)\right)\mathbf{1}_{L>x}\right] \\
= \; & \mathrm{E}\left[\exp\left(-\theta Q+\psi\left(\theta\right)\right)\mathbf{1}_{L>x}\right] \\
\cong \; & \mathrm{E}\left[\exp\left(-\theta Q+\psi\left(\theta\right)\right)\mathbf{1}_{Q>x}\right] \\
\leq \; & \mathrm{E}\left[\exp\left(-\theta x+\psi\left(\theta\right)\right)\mathbf{1}_{Q>x}\right] \text{ si } \theta>0 \\
\leq \; & \exp\left(-\theta x+\psi\left(\theta\right)\right).
\end{aligned}
$$

Notons que la fonction $\exp\left(-\theta x+\psi\left(\theta\right)\right)$ est décroissante en x si $\theta>0$. Par conséquent, le deuxième moment aura tendance à être petit lorsque x est grand (ce qui est le cas pour l'évaluation de la VaR).

Échantillonnage sous la nouvelle mesure I

Échantillonnage stratégique

1. Pour utiliser l'estimateur stratégique

$$\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\theta Q^{(j)}+\psi\left(\theta\right)\right)\mathbf{1}_{L^{(j)}>x},$$

   il faut pouvoir simuler le couple (L, Q) sous la mesure P_θ puisque nous voulons des réplications indépendantes de $\exp\left(-\theta Q+\psi\left(\theta\right)\right)\mathbf{1}_{L>x}$.

2. L et Q ne sont pas simulés directement. Nous pouvons écrire que le couple $(L, Q) = f(\mathbf{Z})$ est une fonction de $\mathbf{Z}$ puisque
   2.1 $\Delta\mathbf{S} = \mathbf{CZ}$ où $\mathbf{Z}$ est $N\left(\mathbf{0}, \mathbf{I}_m\right)$,
   2.2 $L = V\left(\mathbf{S}_0, 0\right) - V\left(\mathbf{S}_0 + \Delta\mathbf{S}, \Delta_t\right)$,
   2.3 $Q = a + \sum_{i=1}^{n}\left(b_i Z_i + \lambda_i Z_i^2\right)$.

Échantillonnage sous la nouvelle mesure II

Échantillonnage stratégique

3. Quelle est la distribution de $\mathbf{Z}$ sous P_θ?
   3.1 Sous P_θ, $\mathbf{Z}$ est normale multivariée où
   3.2 l'espérance de la i ème composante est $\mu_i(\theta) = \frac{\theta b_i}{1+2\theta\lambda_i}$
   3.3 et la matrice de covariance $\Sigma(\theta)$ est diagonale et le $i$ème élément de la diagonale est $\sigma_i^2(\theta) = \frac{1}{1-2\theta\lambda_i}$

Preuve. Rappelons que

$$\begin{aligned}
Q &= a + \sum_{i=1}^{m} \left(b_i Z_i + \lambda_i Z_i^2 \right) \\
\psi(\theta) &= a\theta + \frac{1}{2} \sum_{i=1}^{m} \left[\frac{\theta^2 b_i^2}{1-2\theta\lambda_i} - \ln(1-2\lambda_i\theta) \right] \\
\frac{dP_\theta}{dP} &= \exp(\theta Q - \psi(\theta)) \\
&= \exp\left(\sum_{i=1}^{n} \left(\theta\lambda_i Z_i^2 + \theta b_i Z_i - \frac{1}{2} \frac{\theta^2 b_i^2}{1-2\theta\lambda_i} \right) \right) \left(\prod_{i=1}^{m} \sqrt{1-2\lambda_i\theta} \right)
\end{aligned}$$

Échantillonnage sous la nouvelle mesure III

Échantillonnage stratégique

.

$$
\begin{aligned}
& P_\theta\left(Z_1 \leq z_1, \ldots, Z_m \leq z_m\right) \\
= \; & \mathrm{E}_\theta\left(\mathbf{1}_{Z_1 \leq z_1, \ldots, Z_m \leq z_m}\right) \\
= \; & \mathrm{E}\left(\frac{dP_\theta}{dP} \mathbf{1}_{Z_1 \leq z_1, \ldots, Z_m \leq z_m}\right) \\
= \; & \left(\prod_{i=1}^m \sqrt{1-2\lambda_i\theta}\right) \mathrm{E}\left(e^{\sum_{i=1}^m\left(\theta\lambda_i Z_i^2+\theta b_i Z_i-\frac{1}{2}\frac{\theta^2 b_i^2}{1-2\theta\lambda_i}\right)} \mathbf{1}_{Z_1 \leq z_1, \ldots, Z_m \leq z_m}\right) \\
= \; & \left(\prod_{i=1}^m \sqrt{1-2\lambda_i\theta}\right) \int_{-\infty}^{z_1} \ldots \int_{-\infty}^{z_m} e^{\sum_{i=1}^m\left(\theta\lambda_i z_i^2+\theta b_i z_i-\frac{1}{2}\frac{\theta^2 b_i^2}{1-2\theta\lambda_i}\right)} \frac{e^{-\frac{1}{2}\sum_{i=1}^n z_i^2}}{(2\pi)^{\frac{m}{2}}} dz_1 \ldots dz_m \\
= \; & \int_{-\infty}^{z_1} \ldots \int_{-\infty}^{z_m} \left(\prod_{i=1}^m f\left(z_i, \mu_i(\theta), \sigma_i^2(\theta)\right)\right) dz_1 \ldots dz_m
\end{aligned}
$$

où

Échantillonnage sous la nouvelle mesure IV

Échantillonnage stratégique

- $f\left(z_i, \mu_i(\theta), \sigma_i^2(\theta)\right) = \frac{\sqrt{1-2\lambda_i\theta}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{\left(z_i - \frac{\theta b_i}{1-2\theta\lambda_i}\right)^2}{\frac{1}{(1-2\theta\lambda_i)}}\right)$ est la fonction de densité d'une v.a. de loi normale d'espérance $\mu_i(\theta)$ et de variance $\sigma_i^2(\theta)$ et
- $\prod_{i=1}^{m} f\left(z_i, \mu_i(\theta), \sigma_i^2(\theta)\right)$ est la fonction de densité conjointe de m v.a. gaussiennes indépendantes. ■

Choix du theta I
Échantillonnage stratégique

1. Nous voulons estimer $P(L > x)$.
2. Nous prendrons une décision à partir de $P(Q > x)$.
3. Rappelons que nous n'avons pas la valeur exacte du deuxième moment de notre estimateur mais nous avons une borne :

$$\begin{aligned} & \mathrm{E}_\theta\left[\exp\left(-2\theta Q + 2\psi(\theta)\right)\mathbf{1}_{Q>x}\right] \\ = & \mathrm{E}\left[\exp\left(-\theta Q + \psi(\theta)\right)\mathbf{1}_{Q>x}\right] \\ \leq & \exp\left(-\theta x + \psi(\theta)\right) \text{ si } \theta > 0. \end{aligned}$$

4. La borne du deuxième moment sera minimisée.

Choix du theta II

Échantillonnage stratégique

4.1 Parce que ψ est convexe (c'est une fonction génératrice de cumulants), le minimum est atteint au point θ_x satisfaisant

$$\psi'(\theta_x) = x.$$

Preuve. $\frac{d}{d\theta} \exp\left(-\theta x + \psi(\theta)\right)\Big|_{\theta=\theta_x} = 0 \Leftrightarrow \left(-x + \psi'(\theta_x)\right) \exp\left(-\theta_x x + \psi(\theta_x)\right) = 0 \Leftrightarrow \psi'(\theta_x) = x$. ■

Interprétation

Échantillonnage stratégique

1. Rappelons que $\psi(\theta) = \ln \mathrm{E}\left[e^{\theta Q}\right]$.
2. $$\begin{aligned} \psi^{'}(\theta) &= \frac{d}{d\theta} \ln \mathrm{E}\left[e^{\theta Q}\right] = \frac{\frac{d}{d\theta}\mathrm{E}\left[e^{\theta Q}\right]}{\mathrm{E}\left[e^{\theta Q}\right]} \\ &= \frac{\mathrm{E}\left[Qe^{\theta Q}\right]}{\exp\left(\psi(\theta)\right)} = \mathrm{E}\left[Qe^{\theta Q - \psi(\theta)}\right] = \mathrm{E}_{\theta}\left[Q\right]. \end{aligned}$$
3. Par conséquent, $\mathrm{E}_{\theta_x}\left[Q\right] = \psi^{'}(\theta_x) = x$.
4. En choisissant θ_x, l'échantillonnage stratégique produit des valeurs qui sont autour de x alors qu'avec la densité originale, x se trouve à être dans l'aile droite de la distributon de Q.

Algorithme

Échantillonnage stratégique

1. Simulation des scénarios de pertes
   1.1 Générer $\mathbf{Z}$ suivant $N\left(\mu\left(\theta_x\right), \Sigma\left(\theta_x\right)\right)$.
   1.2 Évaluer $Q = a + \sum_{i=1}^{n}\left(b_i Z_i + \lambda_i Z_i^2\right)$.
   1.3 Poser $\Delta\mathbf{S} = \mathbf{CZ}$.
   1.4 Déterminer la valeur du portefeuille $V\left(\mathbf{S}_0 + \Delta\mathbf{S}, \Delta_t\right)$.
   1.5 Calculer la perte $L = V\left(\mathbf{S}_0, \mathbf{0}\right) - V\left(\mathbf{S}_0 + \Delta\mathbf{S}, \Delta_t\right)$.
   1.6 Calculer $\exp\left(-\theta_x Q + \psi\left(\theta_x\right)\right)\mathbf{1}_{L>x}$
2. Calculer la moyenne échantillonnale

$$\frac{1}{n}\sum_{j=1}^{n}\exp\left(-\theta_x Q^{(j)} + \psi\left(\theta_x\right)\right)\mathbf{1}_{L^{(j)}>x}.$$

   C'est un estimateur de $P\left(L > x\right)$.
3. Calculer les intervalles de confiance et les marges d'erreur selon les techniques habituelles.

Échantillonnage stratifié

Introduction

Échantillonnage stratifié

1. L'échantillonnage stratifié est utilisé en combinaison avec l'échantillonnage stratégique afin d'avoir une réduction de variance supplémentaire.
2. C'est la variable Q qui sera stratifiée.
3. Pour appliquer l'échantillonnage stratifié à l'estimateur $\exp\left(-\theta_x Q + \psi\left(\theta_x\right)\right) \mathbf{1}_{L>x}$, il faut donc
   3.1 définir les strates A_k, $k = 1, ..., K$ pour lesquelles $P_{\theta_x}\left(Q \in A_k\right)$ est connue
   3.2 et déterminer un algorithme permettant de simuler le couple (Q, L) conditionnellement à ce que Q soit dans une strate particulière.

Constructions des strates

Échantillonnage stratifié

1. Rappelons que nous savons comment calculer la distribution de Q sous la mesure originale P :
$$\begin{aligned} & P\left[Q \leq x\right] - P\left[Q \leq x - y\right] \\ = \; & \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left(\phi\left(u\right)\left[\frac{e^{iu\,y} - 1}{iu}\right] e^{-iux}\right) du \end{aligned}$$

2. Cependant, nous devons échantillonner sous P_{θ_x}. Heureusement, Q demeure une fonction quadratique de variables aléatoires gaussiennes sous la nouvelle mesure. En effet,
   2.1 $Q = a + \sum_{i=1}^m \left(b_i Z_i + \lambda_i Z_i^2\right)$ où, sous P, les Z_i sont iid $N\left(0,1\right)$.
   2.2 Sous P_{θ_x} les Z_i sont indépendantes de loi $N\left(\mu_i\left(\theta_x\right), \sigma_i^2\left(\theta_x\right)\right)$.
   2.3 Posons $\widetilde{Z}_i = \frac{Z_i - \mu_i(\theta_x)}{\sigma_i(\theta_x)}$. Sous P_{θ_x}, les $\widetilde{Z}_i$ sont iid $N\left(0,1\right)$.
   2.4
$$\begin{aligned} Q &= a + \sum_{i=1}^m \left(b_i\left(\mu_i\left(\theta_x\right) + \sigma_i\left(\theta_x\right)\widetilde{Z}_i\right) + \lambda_i\left(\mu_i\left(\theta_x\right) + \sigma_i\left(\theta_x\right)\widetilde{Z}_i\right)^2\right) \\ &= \widetilde{a} + \sum_{i=1}^m \left(\widetilde{b}_i \widetilde{Z}_i + \widetilde{\lambda}_i \widetilde{Z}_i^2\right) \end{aligned}$$

Échantillonnage

Échantillonnage stratifié

- Il serait possible de simuler Q à l'aide de la méthode d'inversion. Cependant, la simulation simultanée de Q et L est plus difficile.
- Rappelons que ces variables ne sont pas simulées directement:

1. Générer $\mathbf{Z}$ suivant $N\left(\mu\left(\theta_x\right),\Sigma\left(\theta_x\right)\right)$.
2. Évaluer $Q = a + \sum_{i=1}^{n}\left(b_i Z_i + \lambda_i Z_i^2\right)$.
3. Poser $\Delta\mathbf{S} = \mathbf{CZ}$.
4. Déterminer la valeur du portefeuille $V\left(\mathbf{S}_0 + \Delta\mathbf{S}, \Delta_t\right)$.
5. Calculer la perte $L = V\left(\mathbf{S}_0, \mathbf{0}\right) - V\left(\mathbf{S}_0 + \Delta\mathbf{S}, \Delta_t\right)$.

- Il faudrait pouvoir simuler (Q, L) conditionnellement à ce que $Q \in A_k$.
- Pour cette raison, c'est la méthode du rejet qui est utilisée.