École des Mines de Nancy, année 2011 –12

Méthodes probabilistes pour la simulation [SG241]
Examen final

mardi 10 avril 2012, 8 h 30

PROBLÈME 1 — La tartiflette géante

☛ Dans tout cet exercice, l'énoncé suit un formalisme en vecteurs-*lignes*.

À l'occasion de leur mariage, deux fiancés ont prévu d'organiser une tartiflette géante. La tartiflette est un plat constitué de pommes de terres, de reblochon, de lardons et d'oignons, lesquels sont tous des denrées périssables et donc non stockables. Les fiancés ont ainsi prévu que, dans six mois, il leur faudra acheter 50 kg de pommes de terres, 20 kg de reblochon, 10 kg de lardons et 10 kg d'oignons. À l'heure actuelle, ces quatre ingrédients coûtent respectivement (pour les quantités ci-dessus) 40 €, 240 €, 100 € et 10 €, soit un total de 390 €. Les fiancés ont bloqué un budget prévisionnel de 500 € pour l'achat de ces ingrédients, afin d'avoir de la marge sur la variation possible des prix.

Dans l'éventualité où le prix dépasserait malgré tout 500 €, il leur faudra faire appel à leurs parents... La question est de savoir, en moyenne, combien les fiancés devront demander à leurs parents (ce sera 0 € la plupart du temps, mais il y a quand même une probabilité non nulle que cela soit strictement positif). Autrement dit, si C est le coût total des ingrédients le jour de leur achat (en euros), ils veulent évaluer

$$\mathbb{E}((C-500)_+), \tag{1}$$

où je rappelle que pour $a \in \mathbb{R}$, "a_+" note la partie positive *de a, càd.* $\max\{a, 0\}$.

D'après l'étude des cours du marché, on peut faire la modélisation suivante : on peut prédire le prix des ingrédients respectifs dans six mois comme suivant la loi de $(40e^X, 240e^Y, 100e^Z, 10e^T)$, où (X, Y, Z, T) est un vecteur aléatoire gaussien centré de matrice de covariance

$$M := \begin{pmatrix} 0{,}2 & 0{,}02 & 0{,}02 & 0{,}02 \\ 0{,}02 & 0{,}05 & 0{,}02 & 0{,}02 \\ 0{,}02 & 0{,}02 & 0{,}1 & 0{,}02 \\ 0{,}02 & 0{,}02 & 0{,}02 & 0{,}3 \end{pmatrix}. \tag{2}$$

1. Écrire une fonction MATLAB appelée `cout` qui simule la loi du prix total C des ingrédients. Cette fonction ne prendra aucun argument et reverra le prix total simulé.
Indication : Pour l'ensemble de l'examen, on rappelle que la commande MATLAB `doc` permet d'ouvrir l'aide en ligne du logiciel, et que `doc truc` ouvre l'aide relative à la fonction `truc`.
Indication : Pour simuler (X, Y, Z, T), on pourra utiliser la méthode de Cholesky. Rappelons deux faits à cette fin :

– Si $\vec{U}$ est un vecteur-ligne gaussien standard (càd. à coordonnées indépendantes centrées réduites) et $\boldsymbol{A}$ une matrice dont le nombre de lignes est égal à la taille de $\vec{U}$, alors $\vec{U} \times \boldsymbol{A}$ est un vecteur gaussien centré dont la matrice de covariance est $\boldsymbol{A}^\mathsf{T} \times \boldsymbol{A}$.

– Pour $\boldsymbol{A}$ une matrice carrée positive définie, la fonction de MATLAB `chol(A)` renvoie une matrice carrée $\boldsymbol{B}$ de même dimension que $\boldsymbol{A}$ telle que $\boldsymbol{B}^{\mathsf{T}} \times \boldsymbol{B} = \boldsymbol{A}$.

Indication : Sous MATLAB, e^a s'obtient par la fonction `exp(a)`.

2. Écrire une fonction MATLAB appelée `MCnaif` qui évalue $\mathbb{E}((C-500)_+)$. Cette fonction prendra comme argument le nombre N de simulations demandées et renverra un intervalle de confiance à 2 sigmas pour la quantité évaluée.

On veut maintenant améliorer la précision de la méthode par échantillonnage préférentiel. Notons $\mathbb{P}$ la loi véritable décrite ci-dessus. Nous considérons pour loi d'échantillonnage préférentiel une loi telle que (X,Y,Z,T) soit toujours un vecteur gaussien de matrice de covariance M, mais que ce vecteur ne soit plus centré : l'espérance de ce vecteur sera alors

$$L := (0{,}3; 0{,}3; 0{,}3; 0{,}2). \tag{3}$$

On note $\mathbb{P}'$ la loi correspondante.

3. Exprimer la densité relative de $\mathbb{P}'$ par rapport à $\mathbb{P}$ en fonction de X, Y, Z et T. On écrira juste une formule littérale, càd. qu'on utilisera les notations "M" et "L" sans chercher à les développer selon leurs valeurs (2) et (3).

Indication : Puisque les quantités simulées ne dépendent que de X, Y, Z et T, la densité relative de $\mathbb{P}'$ par rapport à $\mathbb{P}$ est simplement la densité relative de la loi de (X,Y,Z,T) sous $\mathbb{P}'$ par rapport à sa loi sous $\mathbb{P}$.

Indication : On rappelle que si $\vec{V}$ est un vecteur(-ligne) (de dimension k) gaussien de moyenne $\vec{m}$ et de covariance $\boldsymbol{C}$ supposée non dégénérée, alors la densité de ce vecteur par rapport à la mesure de Lebesgue sur $\mathbb{R}^k$ est donnée par

$$\frac{d\mathbb{P}(\vec{V}=\vec{v})}{d\vec{v}} = \frac{1}{\sqrt{2\pi}^k |\mathrm{dét}\,\boldsymbol{C}|^{1/2}} \exp\big(-\tfrac{1}{2}(\vec{v}-\vec{m}) \times \boldsymbol{C}^{-1} \times (\vec{v}-\vec{m})^{\mathsf{T}}\big). \tag{4}$$

4. Grâce au résultat de la question précédente, écrire une fonction MATLAB appelée `MCpref` qui évalue $\mathbb{E}((C-500)_+)$ par la méthode d'échantillonnage préférentiel évoquée ci-dessus. (Si vous n'êtes pas arrivé à faire la question précédente, vous pourrez faire comme si la densité relative était $e^{X-YZ}\sqrt{1+T^2}$ — c'est totalement faux, mais au moins je pourrai voir si vous avez compris le principe). De même que la fonction `Mcnaif`, cette fonction prendra comme unique argument le nombre N de simulations demandées et renverra un intervalle de confiance à 2 sigmas pour la quantité évaluée.

Indication : On rappelle qu'un vecteur gaussien décentré peut s'obtenir comme la somme de son vecteur espérance (déterministe) et d'un vecteur gaussien centré (aléatoire) de même matrice de covariance.

Indication : Sous MATLAB, l'inverse d'une matrice $\boldsymbol{A}$ s'obtient par la fonction `inv(A)`.

5. En modifiant les deux programmes pour faire en sorte de calculer aussi leurs efficacités respectives, comparer les performances de `MCnaif` et `MCpref`.

PROBLÈME 2 — Cours à volatilité variable

Pour commencer, rappelons qu'on peut voir le mouvement brownien géométrique (avec dérive) comme un processus $(A_t)_{t\geqslant 0}$ suivant l'« équation différentielle stochastique » :

$$dA_t = \sigma A_t dW_t + (r + \tfrac{1}{2}\sigma^2) A_t dt. \tag{5}$$

Cela signifie la chose suivante : si on a simulé tout le processus jusqu'au temps t et qu'on souhaite trouver sa valeur en $t+\tau$ où τ est un très petit pas de temps, alors on *peut trouver cette valeur par la formule* (5)*, où dA_t désigne $A_{t+\tau} - A_t$, dt vaut τ et dW_t représente l'accroissement d'un mouvement brownien entre les instants t et $t+\tau$, càd. une variable aléatoire de loi $\mathcal{N}(0,\tau)$.*

Quand on utilise (5) *pour modéliser l'évolution du cours d'un actif financier, t est le temps et s'exprime, par exemple, en années (*a*), A_t est la valeur de l'actif et s'exprime en euros (€), r est plus ou moins assimilable au taux sans risque et s'exprime en* a^{-1}*, et enfin σ est appelé la* volatilité *et s'exprime en* $\mathrm{a}^{-1/2}$.

On considère maintenant le processus stochastique suivant, où l'objet qui évolue n'est pas un simple réel A_t mais un couple *(A_t, β_t) à valeurs dans $\mathbb{R}_+ \times \mathbb{R}$ qui évolue selon l'équation différentielle stochastique suivante :*

$$\left\{\begin{array}{rcl} dA_t & = & e^{\beta_t}\sigma A_t dW_t^{(1)} + (r + \frac{1}{2}e^{2\beta_t}\sigma^2)A_t dt\,; \\ d\beta_t & = & \gamma dW_t^{(2)} - \lambda\beta_t dt, \end{array}\right. \tag{6}$$

$W^{(1)}$ et $W^{(2)}$ étant des mouvement browniens indépendants. Les paramètres de ce modèle sont σ (en $\mathrm{a}^{-1/2}$*), r (en* a^{-1}*), γ (en* $\mathrm{a}^{-1/2}$*) et λ (en* a^{-1}*).*

6. Écrire une fonction `MATLAB` appelée `simulation` qui simule le processus $(A_t)_t$ sur l'intervalle de temps $[0, T]$. Cette fonction prendra les arguments suivants : donnée de σ, de r, de γ, de λ, valeur initiale A_0 pour A_t, valeur initiale β_0 pour β_t, choix de T, et choix du nombre M de sous-intervalles de temps en combien on divise $[0, T]$ pour la simulation. Cette fonction aura comme résultat de dessiner l'évolution de $(A_t)_{t\geqslant 0}$ — et seulement de $(A_t)_t$ — sur l'intervalle de temps considéré.
Indication : On rappelle que la syntaxe de la fonction `MATLAB` `plot` est la suivante : si `X` et `Y` sont deux vecteurs-lignes de même taille z, `plot(X,Y)` trace la courbe bidimensionnelle paramétrée $\big(X(i), Y(i)\big)_{1\leqslant i\leqslant z}$, en reliant les différents points par des segments de droite.
Indication : Pour tester votre fonction, voici un jeu de paramètres pertinent : $\sigma = 0{,}2$, $r = 0{,}1$, $\gamma = 4$, $\lambda = 4$, $A_0 = 20$, $\beta_0 = 0$, $T = 1$, $M = 730$.

7. Comment comprenez-vous le modèle (6) ? Quel sens donneriez-vous aux différentes quantités ?
Indication : Le titre de ce problème donne déjà une indication...
Indication : On pourra d'abord essayer de comprendre la signification du processus $(\beta_t)_{t\geqslant 0}$, observant que son évolution ne dépend pas de ce qui se passe pour A_t.

Si on procède au changement de variables $\Lambda_t := \ln(A_t)$, nous admettrons qu'on peut faire comme si l'équation (6) *devenait la suivante :*

$$\left\{\begin{array}{rcl} d\Lambda_t & = & e^{\beta_t}\sigma dW_t^{(1)} + r dt\,; \\ d\beta_t & = & \gamma dW_t^{(2)} - \lambda\beta_t dt. \end{array}\right. \tag{7}$$

8. Conditionnellement à la donnée de Λ_0 et de toute la trajectoire de $(\beta_t)_{0\leqslant t\leqslant T}$, montrer que Λ_T suit une loi normale dont on exprimera les paramètres en fonction des paramètres du modèle, de Λ_0 et de toute la trajectoire $(\beta_t)_{0\leqslant t\leqslant T}$. (On s'autorisera à considérer les pas de temps comme infinitésimalement petits, de façon à tronquer les développements limités ou à assimiler sommes et intégrales).

9. Montrer que si X est une variable réelle suivant la loi normale décentrée $\mathcal{N}(m, \sigma^2)$, alors $\mathbb{E}(e^X) = \exp(m + \frac{1}{2}\sigma^2)$.

La question qu'on se pose maintenant est celle d'un acheteur qui sait qu'il aura besoin d'acheter l'actif A à l'instant T et qui souhaite dès à présent ($t = 0$) s'assu-

rer contre les variations possibles de son prix d'ici-là. Bref, notre acheteur a besoin d'évaluer $\mathbb{E}(A_T)$*,* [*] *sachant que pour l'instant il ne connaît que* A_0 *et* β_0*.*

10. Conditionnellement à la donnée de A_0 et de l'ensemble de la trajectoire de $(\beta_t)_{0\leqslant t\leqslant T}$, quelle est l'espérance de A_T ?
Indication : On pourra utiliser les résultats des deux questions précédentes.

11. Grâce à la question précédente, écrire un algorithme `MCcondi` qui utilise la technique de conditionnement pour évaluer $\mathbb{E}(A_T)$ (A_0 et β_0 étant déjà donnés), le conditionnement se faisant par rapport à l'ensemble de la trajectoire $(\beta_t)_{0\leqslant t\leqslant T}$. Les paramètres du modèle sont fixés à $\sigma = 0{,}2$, $r = 0{,}1$, $\gamma = 4$ et $\lambda = 4$. La fonction prendra comme paramètres les valeurs de A_0 et β_0, la valeur de l'échéance T, le nombre M de sous-intervalles de temps pour la simulation et le nombre N se simulations effectuées pour la méthode de Monte-Carlo, et renverra un intervalle de confiance à 2 sigmas pour $\mathbb{E}(A_T)$.

PROBLÈME 3 — La théorie et la pratique

Soit

$$I := \int_{-\infty}^{+\infty} \frac{1}{1+x^6}\,dx. \tag{8}$$

12. Par la méthode de votre choix (pas forcément Monte-Carlo, donc), évaluer numériquement I avec précision (6 décimales exactes au moins). On pourra également calculer une expression analytique de I.
Indication : En entrant la commande `format long` dans la console de MATLAB, on bascule la convention d'affichage des nombres réels de sorte à ce qu'ils soient donnés avec une précision de 15 décimales. Pour revenir à la convention d'affichage initiale, on utilise la commande `format short`.

On voudrait maintenant évaluer I *par la méthode de Monte-Carlo, avec un intervalle de confiance à 90 %. La loi d'échantillonnage sera la loi double-exponentielle de paramètre* 1*, définie par*

$$d\mathbb{P}(x) := \tfrac{1}{2}e^{-|x|}dx. \tag{9}$$

13. Indépendamment du fait que la méthode de Monte-Carlo soit peu indiquée pour ce genre de situation, qu'est-ce qui ne va pas du point de vue théorique ?

14. Quels peuvent être les effets du problème théorique évoqué à la question précédente ? Il y a deux items attendus pour la réponse.
Indication : Fiabilité, efficacité...

15. Programmer la méthode de Monte-Carlo et tester les effets évoqués à la question précédente. Commenter.

16. Si on mène la méthode de Monte-Carlo avec N simulations, justifier de l'existence d'un $K < \infty$ (dépendant de N) tel qu'on puisse considérer, dans une certaine mesure, que l'intégrale que calcule notre méthode de Monte-Carlo est en fait

$$I(K) := \int_{-K}^{+K} \frac{1}{1+x^6}\,dx. \tag{10}$$

Proposer une valeur pour K en fonction de N.

[*]. J'ai écrit que r était plus ou moins assimilable au taux sans risque, mais ici on ne s'intéressera pas à la possibilité de placer son argent sans risque : c'est bien $\mathbb{E}(A_T)$ qu'on veut calculer et pas $\mathbb{E}(e^{-rT}A_T)$.

17. Notons $\sigma(K)$ l'écart-type de la variable aléatoire qui apparaît quand on évalue $I(K)$. Tracer (à la main ou à la machine) l'allure de σ en fonction de K sur l'intervalle $[3, 48]$.

18. À l'aide du résultat des deux questions précédentes, donner une explication théorique au phénomène constaté à la question 15.