Méthode de Monte-Carlo & Application aux Processus Aléatoires

# Devoir à la maison en vue de la seconde session d'examen

É. N. S. des Mines de Nancy / Génie Industriel & Mathématiques Appliquées

31 mars 2016

## Problème I – Placement prudent

*Suite à un excellent exercice budgétaire, une entreprise se retrouve avec une quantité importante $Q_0 := 1\,000$ k€ de liquidités en caisse. Comme il n'y a aucun investissement industriel intéressant à faire à court terme, le conseil d'administration demande à ce que cet argent soit placé en bourse pour l'année à venir (durée notée $T$). Néanmoins, il est précisé que le placement doit être fait en sorte que l'entreprise soit à tout moment en mesure de récupérer en caisse une quantité plancher $P := 600$ k€ d'argent en liquidant immédiatement ses positions.*

*Bien entendu, une solution simple consisterait à placer la quantité $P$ dans les coffres, et à investir en bourse le montant $(Q_0 - P)$ restant. Néanmoins le directeur financier décide de faire quelque chose d'un peu plus subtil : la quantité d'argent placée en bourse sera ajustée en fonction de l'évolution des cours de manière à ce que, à tout instant, si on devait liquider les positions en bradant l'actif à seulement $70\ \% =: (1-\lambda)$ de sa valeur de marché, on récupère toujours la quantité $P$ (ou plus si on est trop riche pour faire autrement). (Précisons que l'entreprise n'empruntera pas d'argent ni ne vendra d'actions à découvert, c.-à-d. que les quantités d'argent dans le coffre et d'actions détenues seront toujours positives ou nulles).*

*On suppose que le placement en bourse se fait sur un unique actif (ou panier d'actifs) dont le montant à l'instant $t$ est noté $A_t$, et que le processus $A$ suit l'équation différentielle stochastique suivante :*

$$dA_t = rA_t dt + \sigma A_t dW_t,$$

*avec $r = 0{,}10\ \mathrm{an}^{-1}$ et $\sigma = 0{,}50\ \mathrm{an}^{-1/2}$, $W_t$ désignant comme d'habitude un mouvement brownien standard, avec la condition initiale $A_0 = 100$ €.*

**Question 1 –** Notons $X_t$ la quantité de titres de l'actif $A$ possédée à l'instant $t$, resp. $R_t$ la quantité d'argent conservée dans les coffres à l'instant $t$, et $Q_t$ la quantité d'argent que cela représenterait si on liquidait immédiatement les positions au cours du marché. Justifier que, à tout instant,

$$X_t = \min\big(Q_t, (Q_t - P) \,/\, \lambda\big) \,/\, A_t,$$

et que l'évolution de $Q$ est régie par l'équation différentielle stochastique

$$dQ_t = X_t dA_t.$$

**Question 2 –** Les évolutions des processus stochastiques suivants sont-elles markoviennes : $A$, $Q$, $XA$, $X$ ? Justifier.

**Question 3 –** Écrire une fonction `Q3` qui simule, sur l'intervalle de temps $[0, T]$ (0 désignant l'instant actuel), l'évolution des quantités $X_0 A_t, X_t A_t, R_t, Q_t$ (sur la même échelle, en couleurs resp. noir, rouge, vert et bleu). L'axe des abscisses sera gradué en années et celui des ordonnées en kiloeuros. On utilisera 5 000 pas de temps pour la simulation. La fonction ne prendra aucun argument, affichera le graphique, et ne renverra aucune valeur.

**Question 4 –** Écrire une fonction `Q4` qui estime, par la méthode de Monte-Carlo, l'espérance de $Q_T$, assortie d'un intervalle de confiance à 80 %. La fonction prendra en argument le nombre de simulations demandées, affichera le résultat ainsi que le temps pris par le calcul, et ne renverra aucune valeur.

**Question 5 –** Justifier que le processus stochastique $e^{-rt} A_t$ (indexé par le temps $t$) est une martingale. En déduire qu'on a

$$\mathbb{E}(A_T) = e^{rT} A_0.$$

**Question 6 –** Expliquer pourquoi le choix de la variable de contrôle $X_0 A_T$ est (au moins à première vue) pertinent pour affiner la précision de l'estimation de $\mathbb{E}(Q_T)$.

**Question 7 –** Écrire une fonction `Q7` implémentant la méthode de la variable de contrôle (avec la variable de contrôle $X_0 A_T$) pour estimer $\mathbb{E}(Q_T)$. La fonction prendra en argument le nombre de simulations demandées, affichera le résultat (et son intervalle de confiance) ainsi que le temps pris par le calcul, et ne renverra aucune valeur. Gagne-t-on en vitesse de calcul par rapport à la question 4 ?

*Une ingénieure de la direction financière fait remarquer que l'espérance de $Q_T$ n'est peut-être pas le meilleur indicateur de la performance du placement, « car (explique-t-elle), avoir 10 % de chances de finir avec 20 000 k€ en caisse et 90 % de risques de finir avec 100 k€ serait certainement une perspective moins attirante qu'avoir 50 % de chances de finir avec 2 000 M€ et 50 % de chances de finir avec 1 000 M€... ». Selon elle, ce qui compte, c'est l'*espérance géométrique *de $Q_T$ et non pas son espérance* stricto sensu*, l'espérance géométrique $\tilde{\mathbb{E}}(X)$ d'une variable aléatoire réelle strictement positive $X$ étant définie par la relation :*

$$\ln(\tilde{\mathbb{E}}(X)) \coloneqq \mathbb{E}(\ln(X)).$$

**Question 8 –** Écrire une fonction `Q8` estimant, par la méthode de Monte-Carlo, l'espérance géométrique de $Q_T$, assortie d'un intervalle de confiance à 80 %. (On ne demande pas d'implémenter de variable de contrôle dans cette question). La fonction prendra en argument le nombre de simulations demandées, affichera le résultat, et ne renverra aucune valeur.

## Problème II – Ping-pong

*Dans ce problème, nous chercherons dans un premier temps à simuler un match de ping-pong entre deux joueurs $A$ et $B$ ; et par la suite nous utiliserons la méthode de Monte-Carlo pour estimer le nombre moyen $n$ de points joués durant la rencontre ainsi que la probabilité $p_A$ que le joueur $A$ gagne.*

*Rappelons que le* ping-pong *(également appelé « tennis de table ») est un sport de raquettes dans lequel les joueurs s'affrontent en 3 ou 4 manches gagnantes* [∗]. *Pour gagner une manche, il faut arriver à avoir marqué au moins 11 points (dans la manche)* et *à avoir marqué (toujours dans la manche) au moins 2 points de plus que son adversaire* [†].

[∗]. Lorsqu'on dit qu'un match se déroule « en $g$ manches gagnantes », cela signifie que le premier des deux joueurs à remporter $g$ manches sera déclaré vainqueur du match. Le match se déroulera par conséquent en un maximum de $(2g-1)$ manches — c'est pourquoi on parle aussi de match « au meilleur des $(2g-1)$ manches ».

[†]. Ainsi, si le joueur $A$ mène 10-6 ou 11-10, ce n'est pas suffisant pour remporter la manche ; en revanche il pourra remporter la manche sur un score de 11-8 ou de 13-11.

*À chaque point joué, il faut savoir qui donne le premier coup de raquette (coup appelé* service*). La règle qui décide à qui c'est de servir est la suivante :*

— *Dans la première manche du match, le joueur qui sert le premier point est tiré au sort équitablement à pile ou face.*
— *Dans les manches suivantes, le joueur qui sert le premier point de la manche alterne à chaque manche : par exemple, si $A$ sert le premier point de la 1re manche, c'est $B$ qui servivra le premier point de la 2e manche, puis $A$ qui servira le premier point de la 3e manche, etc.*
— *À l'intérieur d'une même manche, le service alterne tous les 2 points disputés pendant les 20 premiers points ; mais si le score est de 10-10 à l'issue des 20 premiers points, à partir de ce moment-là le service alternera à chaque point disputé jusqu'à la fin de la manche. Par exemple, si $A$ sert le 1er point de la manche, alors il servira aussi le 2e point de la manche, puis $B$ servira les 3e et 4e points, puis $A$ servira les 5e et 6e, et ainsi de suite ; et si le score est de 10-10 à l'issue de 20 points (après donc que $B$ aura servi les 19e et 20e points), $A$ servira le 21e point de la manche, puis $B$ le 22e, $A$ le 23e, etc.*

*Dans les simulations que l'on considèrera pour la suite, nous ferons les hypothèses suivantes :*

— *Tout au long de la partie, le joueur $A$ a la probabilité $p_{\mathrm{S}}$ ('S' pour « service ») de gagner le point lorsque c'est lui qui sert, et une probabilité $p_{\mathrm{R}}$ ('R' pour « retour ») de gagner le point lorsque c'est $B$ qui sert.*
— *Les résultats de chaque point sont indépendants : savoir qui a gagné les points précédents n'a aucune incidence sur qui gagnera les points à venir (si ce n'est à travers l'identité du joueur qui va servir, évidemment).*

**Question 9 –** Dans cette question, on cherche à simuler l'issue d'*une manche* entre $A$ et $B$, sachant qui sert en premier. Écrire une fonction `Q9` dont les arguments d'entrée sont $(p_{\mathrm{S}}, p_{\mathrm{R}}, s)$, où $s$ désigne un booléen valant « vrai » (ou un entier valant 1) si c'est $A$ qui sert en premier dans la manche simulée, resp. « faux » (ou 0) si c'est $B$ qui sert en premier. La fonction renverra, sous forme d'un 2-vecteur ligne, le couple $(N, G)$, où $N$ désigne le nombre total de points disputés dans la manche simulée, et $G$ désigne un booléen valant « vrai » (ou un entier valant 1) si $A$ a gagné la manche, resp. « faux » (ou 0) si c'est $B$ qui a gagné.

**Question 10 –** Dans cette question,on cherche à simuler l'issue d'*un match* en $k$ manches gagnantes entre $A$ et $B$. Écrire une fonction `Q10` (en vous servant de la fonction `Q9`) dont les arguments d'entrée sont $(p_{\mathrm{S}}, p_{\mathrm{R}}, k)$, où $k$ désigne le nombre de manches gagnantes. La fonction renverra, sous forme d'un 2-vecteur ligne, le couple $(N, G)$, où $N$ désigne le nombre total de points disputés dans le match simulé, et $G$ désigne un booléen valant « vrai » (ou un entier valant 1) si $A$ a gagné le match, resp. « faux » (ou 0) si c'est $B$ qui a gagné.

☛ *Dans la suite, quand il sera demandé un intervalle de confiance, on prendra un intervalle de confiance à 95 %.*

**Question 11 –** Dans cette question, on veut employer la méthode de Monte-Carlo pour estimer le nombre moyen de points $n$ que dure un match en $k$ manches gagnantes entre $A$ et $B$, ainsi que la probabilité $p_A$ pour que $A$ gagne. Écrire une fonction `Q11` dont les arguments d'entrée sont $(p_{\mathrm{S}}, p_{\mathrm{R}}, k, K)$, où $K$ désigne le nombre de simulations pour la méthode de Monte-Carlo, qui estime la durée moyenne (en nombre de points) $n$ d'un match en $k$ manches gagnantes entre $A$ et $B$, ainsi que la probabilité $p_A$ que $A$ gagne le match. La fonction affichera les intervalles de confiance de $n$ et $p_A$, et ne renverra aucune valeur.

**Question 12 –** Constater expérimentalement, en appliquant votre fonction `Q11` pour différentes valeurs de $(p_\mathrm{S}, p_\mathrm{R})$, que la joueur $A$ peut être considéré comme plus fort que le joueur $B$ si $p_\mathrm{S} + p_\mathrm{R} > 1$, resp. moins fort si $p_\mathrm{S} + p_\mathrm{R} < 1$, resp. aussi fort si $p_\mathrm{S} + p_\mathrm{R} = 1$. Expliquer par un argument théorique rigoureux [‡] pourquoi cela est effectivement vrai.

*On considère maintenant l'organisateur d'un tournoi de ping-pong, qui est soumis à deux contraintes distinctes : d'une part, il souhaite que son tournoi récompense effectivement le meilleur joueur ; et d'autre part, il souhaite que le tournoi se termine dans le délai imparti. Le seul paramètre sur lequel l'organisateur peut influer est le choix du nombre de manches gagnantes (3 ou 4) pour le tournoi. Pour étudier l'impact d'un tel choix, notre organisateur souhaite comparer ce que donnera, en moyenne, l'issue d'un match en 3 manches gagnantes, resp. en 4 manches gagnantes.*

☛ *Dans la question suivante, les nombres de manches gagnantes ne sont pas nécessairement égaux à 3 ou 4.*

**Question 13 –** Soient $k_1 < k_2$. Expliquer, par un argument théorique suffisamment rigoureux, pourquoi un match en $k_2$ manches gagnantes, par rapport à un match en $k_1$ manches gagnantes, aura une plus grande probabilité de désigner effectivement comme vainqueur le plus fort des deux joueurs, mais que le match sera par contre plus long en moyenne.

**Question 14 –** Dans cette question, on va estimer par une méthode de couplage les différences entre les propriétés statistiques $(n, p_A)$ d'un match selon qu'il se dispute en 3 manches gagnantes ou en 4 manches gagnantes. Écrire une fonction `Q14` dont les arguments d'entrée sont $(p_\mathrm{S}, p_\mathrm{R}, K)$, qui estime, par la méthode de Monte-Carlo avec technique de couplage (avec $K$ simulations), la différence $dn$ entre la moyenne du nombre de points disputés selon que le matche se dispute en 3 ou 4 manches gagnantes, et la différence $dp_A$ entre les probabilités que le joueur $A$ a de gagner le match selon les deux situations. La fonction affichera les intervalles de confiance de $dn$ et de $dp_A$, et ne renverra aucune valeur.

*Oublions maintenant notre organisateur. La joueuse $A$, classée 21e mondiale, doit affronter la championne olympique Li* XIAOXIA *(joueuse $B$) dans un match en 3 manches gagnantes. Les bookmakers [§] britanniques cherchent à établir la cote à attribuer à l'éventualité d'une victoire de la joueuse $A$. Via diverses méthodes statistiques, ils estiment que pour ce match, on aura $p_\mathrm{S} = 0{,}45$ et $p_\mathrm{R} = 0{,}35$. Nos bookmakers souhaitent estimer, à partir de cela, la probabilité que la joueuse $A$ gagne.*

**Question 15 –** Donner (en utilisant la fonction `Q11`) un intervalle de confiance pour la probabilité que la 21e mondiale gagne contre la championne olympique, calculé à partir de 1 000 simulations.

*Afin de réduire la largeur de l'intervalle de confiance, nos bookmakers souhaitent utiliser une méthode de Monte-Carlo utilisant un échantillonnage préférentiel. Pour cela, ils vont considérer la loi d'échantillonnage préférentiel dont les paramètres seraient $p'_\mathrm{S} := 0{,}55$ et $p'_\mathrm{R} := 0{,}45$.*

**Question 16 –** Notons $\mathbb{P}$ la loi d'un match avec $p_\mathrm{S} = 0{,}45$ et $p_\mathrm{R} = 0{,}35$, resp. $\mathbb{P}'$ la loi d'un match avec $p'_\mathrm{S} = 0{,}50$ et $p'_\mathrm{R} = 0{,}50$. Montrer que la densité de $\mathbb{P}'$ par rapport à $\mathbb{P}$, pour une

[‡]. Par « rigoureux », je n'entends pas forcément que ce soit une démonstration mathématique complète, mais qu'il y ait au moins un vrai raisonnement convaincant, et pas juste une vague évocation de quelques concepts.

[§]. Preneurs de paris.

éventualité donnée, est égale à

$$\left(\frac{p'_{\mathrm{S}}}{p_{\mathrm{S}}}\right)^{N_A^{\mathrm{S}}}\left(\frac{1-p'_{\mathrm{S}}}{1-p_{\mathrm{S}}}\right)^{N^{\mathrm{S}}-N_A^{\mathrm{S}}}\left(\frac{p'_{\mathrm{R}}}{p_{\mathrm{R}}}\right)^{N_A^{\mathrm{R}}}\left(\frac{1-p'_{\mathrm{R}}}{1-p_{\mathrm{R}}}\right)^{N^{\mathrm{R}}-N_A^{\mathrm{R}}},$$

où $N_A^{\mathrm{S}}$ désigne le nombre de points marqués par $A$ *sur son propre service* au cours de l'ensemble du match, $N_A^{\mathrm{R}}$ désigne le nombre de points marqués par $A$ *sur le service adverse* au cours de l'ensemble du match, $N^{\mathrm{S}}$ désigne le nombre total de points servis par $A$, et $N^{\mathrm{R}}$ désigne le nombre total de points servis par $B$.

**Question 17 –** Grâce au résultat de la question précédente, écrire une fonction `Q17` qui calcule un intervalle de confiance pour la probabilité que la 21[e] mondiale gagne contre la championne olympique, calculé à partir de 1 000 simulations *en utilisant la méthode d'échantillonage préférentiel*. La fonction ne prendra aucun argument en entrée, affichera l'intervalle de confiance, et ne renverra aucune valeur.