École Nationale Supérieure des Mines de Nancy
Département « Génie Industriel & Mathématiques Appliquées »
Année 2014-2015

Contrôle de compétences sur le module « Méthode de Monte-Carlo & Application aux Processus Aléatoires »

durée : $3\ \mathrm{h}$

- L'ensemble des questions vaut $161\ \mathrm{pt}$: $96\ \mathrm{pt}$ pour le problème I et $65\ \mathrm{pt}$ pour le problème II.
- Un score de $55\ \mathrm{pt}$ permettra la validation du module à coup sûr.
- La note 'B' devrait se jouer autour de $77\ \mathrm{pt}$.

DON'T PANIC.

The Hitchhiker's Guide to the Galaxy

PROBLÈME I — *Éclipses*

Une éclipse de Soleil *se produit lorsque, à la nouvelle-lune, la Lune se trouve alignée avec le Soleil du point de vue de l'observateur. Ce problème se propose de déterminer la fréquence de ce phénomène, à partir d'une modélisation légèrement simplifiée.*

Notre modélisation implique les notations et formules données par le tableau 1, ainsi que la formule (1) *ci-dessous :*

$$\begin{aligned} x_{\mathrm{T}} = R_{\mathrm{T}} \times \big(& -\cos\Phi_{\mathrm{T}}\cos\lambda_{\mathrm{T}} \\ &+ \sin E_{\mathrm{T}}\sin\lambda_{\mathrm{S}}\sin\Phi_{\mathrm{T}} + (1-\cos E_{\mathrm{T}})\cos\Phi_{\mathrm{T}}\sin\lambda_{\mathrm{S}}\sin(\lambda_{\mathrm{T}}+\lambda_{\mathrm{S}})\big) \end{aligned} \tag{1}$$

Comme vous le voyez, certaines des variables du modèle sont aléatoires : en effet, dans ce problème nous considérons une nouvelle-lune prise au hasard (uniformément). Les variables aléatoires fondamentales, à savoir $d_{\mathrm{L}}, d_{\mathrm{S}}, \lambda_{\mathrm{S}}, \lambda_{\mathrm{T}}, \varphi_{\mathrm{L}}$*, sont supposées indépendantes dans cette modélisation.* $Beta$ *désigne (des réalisations de) la loi bêta de paramètres* $(1/2, 1/2)$*, que nous définissons maintenant.*

La loi bêta (de paramètres $(1/2, 1/2)$*) est une loi sur l'intervalle* $(0, 1)$*, à densité par rapport à la mesure de Lebesgue, dont la distribution est caractérisée par*

$$\frac{d\mathbb{P}(Beta = x)}{dx} = \mathbf{1}_{x\in(0,1)} \frac{1}{\pi} x^{-1/2}(1-x)^{-1/2}.$$

[1]. Il s'agit de l'angle qu'occupe le rayon de la Lune dans le champ visuel de l'observateur.

[2]. Comme vous le savez, l'orbite lunaire n'est pas parfaitement circulaire ; la distance de la Terre à la Lune oscille donc à chaque orbite entre des valeurs minimale et maximale.

[3]. Le rayon de la Terre n'étant pas négligeable devant la distance Terre – Lune, la distance entre la Lune et l'observateur ne sera pas exactement la même que la distance avec le centre de la Terre...

[4]. Comme vous le savez, l'axe de rotation de la Terre est légèrement incliné par rapport à son orbite autour du Soleil, ce qui cause les saisons. L'obliquité désigne l'amplitude de cette inclinaison.

[5]. Cette quantité vaut 0° à l'équinoxe de printemps, 90° au solstice d'été, 180° à l'équinoxe d'automne et 270° au solstice d'hiver.

[6]. Cette quantité vaut 0° à minuit solaire, et 180° à midi solaire.

[7]. Cette quantité indique en substance à quel point la Lune est plus haut (ou moins haut) que le Soleil dans le ciel.

[8]. Le plan de l'orbite lunaire autour de la Terre ne coïncide pas exactement avec le plan de l'orbite terrestre autour du Soleil ; l'inclinaison désigne l'angle entre ces deux plans.

[9]. J'ai pris la latitude de Nancy ☺

[10]. L'axe des abscisses pointe de la Terre vers le Soleil.

Nom	Signification	Valeur
α_{L}	Rayon angulaire de la Lune [1]	$R_{\mathrm{L}} \,/\, d'_{\mathrm{L}}$
α_{S}	Rayon angulaire du Soleil	$R_{\mathrm{S}} \,/\, d_{\mathrm{S}}$
D_{L}^-	Distance Terre – Lune minimale [2]	$3{,}631 \times 10^8$ m
D_{L}^+	Distance Terre – Lune maximale	$4{,}057 \times 10^8$ m
D_{S}^-	Distance Terre – Soleil minimale	$1{,}471 \times 10^{11}$ m
D_{S}^+	Distance Terre – Soleil maximale	$1{,}521 \times 10^{11}$ m
d_{L}	Distance Terre – Lune	$D_{\mathrm{L}}^- + (D_{\mathrm{L}}^+ - D_{\mathrm{L}}^-)Beta$
d_{S}	Distance Terre – Soleil	$D_{\mathrm{S}}^- + (D_{\mathrm{S}}^+ - D_{\mathrm{S}}^-)Beta$
d'_{L}	Distance de l'observateur à la Lune [3]	$d_{\mathrm{L}} - x_{\mathrm{T}}$
E_{T}	Obliquité de l'axe terrestre vs. l'écliptique [4]	23,44°
λ_{S}	Saison [5]	$Uniforme([0°, 360°])$
λ_{T}	Heure solaire de l'observateur [6]	$Uniforme([0°, 360°])$
φ_{L}	Position de la Lune vs. l'écliptique [7]	$(2Beta - 1)\,\Phi_L$
Φ_{L}	Inclinaison de l'orbite lunaire vs. l'écliptique [8]	5,145°
Φ_{T}	Latitude de l'observateur [9]	48,67°
R_{L}	Rayon de la Lune	$1{,}736 \times 10^6$ m
R_{S}	Rayon du Soleil	$6{,}963 \times 10^8$ m
R_{T}	Rayon de la Terre	$6{,}378 \times 10^6$ m
T_{L}	Période entre deux nouvelles-lunes	29,53 j
T_{S}	Durée d'une année	365,2 j
x_{T}	Abscisse de l'observateur vs. la Terre [10]	*voir formule* (1)

TABLEAU 1 – Variables utlisées dans le problème I

	$\varphi_L + \alpha_L < -\alpha_S$	$\varphi_L + \alpha_L \in [-\alpha_S, \alpha_S]$	$\varphi_L + \alpha_L > \alpha_S$
$\varphi_L - \alpha_L < -\alpha_S$	0	1	2
$\varphi_L - \alpha_L \in [-\alpha_S, \alpha_S]$	*impossible*	1 [11]	1
$\varphi_L - \alpha_L > \alpha_S$	*impossible*	*impossible*	0

TABLEAU 2 – Quel type d'éclipse on a dans quel cas

10 pt **1.** Démontrer que, si U suit une loi uniforme sur $(0, 1)$, alors

$$\frac{1 - \cos(\pi U)}{2}$$

suit la loi bêta.

La présence ou l'absence d'une éclipse dépend alors des valeurs de φ_L, α_L et α_S. Il y a en fait trois situations possibles, auxquelles nous attribuons des numéros de code arbitraires :
- *Pas d'éclipse (0) ;*
- *Éclipse* partielle *(1) ;*
- *Éclipse* totale *(2).*

Le tableau 2 décrit quand on est dans quel cas.

18 pt **2.** Écrire une fonction `simul_eclipse` qui ne prend aucun argument, simule les paramètres d'une nouvelle-lune prise au hasard, et renvoie le numéro de code indiquant s'il y a une éclipse et si oui de quel type.

10 pt **3.** Écrire une fonction `proba_eclipse` qui prend en argument un paramètre `type` dans $\{1, 2\}$ ainsi qu'un nombre de simulations `Nsim`, et qui estime la probabilité qu'une nouvelle-lune prise au hasard corresponde à une éclipse de type (dont le code vaut) au moins $type$, en réalisant N_{sim} simulations de Monte-Carlo [12]. On précisera un intervalle de confiance à 95 % pour cette estimation. Cette fonction se contentera d'afficher son résultat à l'écran, sans rien renvoyer.

Indication — Pour un affichage rudimentaire [13] de la valeur numérique d'une variable `lavariable`, on peut utiliser l'instruction « `disp (lavariable)` ». De même, on peut afficher un texte par « `disp ('le texte à afficher')` ».

8 pt **4.** Justifier que la probabilité que φ_L appartienne à l'intervalle $[\varphi_-, \varphi_+]$ (où on suppose $-\Phi_L \leqslant \varphi_- < \varphi_+ \leqslant \Phi_L$) vaut exactement

$$\frac{1}{\pi}\left(\arcsin\frac{\varphi_+}{\Phi_L} - \arcsin\frac{\varphi_-}{\Phi_L}\right)$$

16 pt **5.** À l'aide de la question précédente, améliorer la fonction `proba_eclipse` en une fonction `proba_eclipse_cond` utilisant une technique de conditionnement pour réduire la variance, le conditionnement étant opéré par rapport à toutes les variables sauf φ_L.

10 pt **6.** Si l'intervalle de confiance pour la probabilité d'une éclipse de type au

[11]. Ce type particulier d'éclipse partielle est qualifiée d'*annulaire*, car la Lune est alors entièrement située devant le Soleil, quoique sans le recouvrir complètement, laissant ainsi subsister un anneau de lumière.

[12]. Par exemple, « `proba-eclipse (1, 10000)` » évaluera la probabilité d'avoir une éclipse partielle ou totale, avec 10 000 simulations.

[13]. C.-à-d. que l'affichage se fera sur une ligne à part — ce qui n'est pas très esthétique, mais sans importance ici.

moins `type` est $p \pm \varepsilon$, avec $\varepsilon \ll p$, justifier que l'intervalle de confiance du temps moyen entre deux éclipses de type au moins `type` est

$$\frac{T_{\mathrm{L}}}{p} \pm \frac{T_{\mathrm{L}}\varepsilon}{p^2}.$$

7 pt **7.** Empiriquement, les relevés astronomiques montrent qu'on a l'occasion d'observer une éclipse (partielle ou totale) une fois tous les $2{,}5\ \mathrm{an}$ environ. Pourtant, si vous avez bien fait vos calculs, vous avez dû trouver un délai à peu près deux fois plus court... Quel phénomène essentiel explique cette différence, à votre avis ?

17 pt **8.** Observe-t-on plus d'éclipses totales de Soleil à Quito (ville située sur l'équateur) ou à Nancy ?

PROBLÈME II — *C'est chaud... !*

Dans ce problème, on étudie l'échauffement d'un circuit électrique au cours d'un certain processus industriel. Le risque qu'on cherche à quantifier est celui de la fonte du circuit, qui aurait des conséquences graves.

On note θ_t la température du circuit à l'instant t. L'évolution de θ_t est donnée par l'équation différentielle stochastique suivante :

$$d\theta_t = \sigma\, dW_t - \lambda\big(\theta_t - \Theta_t\big)\, dt, \tag{2}$$

où W_t désigne un mouvement brownien standard, λ, σ sont des paramètres du modèle (voir le tableau 3 pour les valeurs de ces paramètres), et où $(\Theta_t)_t$ est une fonction déterministe définie par :

$$\Theta_t := \Theta_* - \alpha(t - T_2)^2/2, \tag{3}$$

Θ_*, T_2, α étant d'autres paramètres du modèle (voir encore tableau 3). On a en outre la condition initiale :

$$\theta_{t=0} \sim \mathcal{N}(\Theta_* - \alpha T_2^2/2 - \alpha T_2 \,/\, \lambda, \sigma^2 \,/\, 2\lambda). \tag{4}$$

6 pt **9.** Si la température se mesure en K et que le temps se mesure en min, déduire des équations (2) et (3) dans quelles unités se mesurent les paramètres $\Theta_*, T_2, \alpha, \lambda, \sigma$; puis vérifier l'homogénéité physique de (4).

Paramètre	Valeur
α	0,2
Δ	5
λ	0,2
σ	2,2
T_1	25
T_2	30
T_3	35
T_4	60
Θ_*	1 348
Θ_{f}	1 358

TABLEAU 3 – Valeurs des paramètres du problème II

9 pt **10.** Simuler le processus $(\theta_t)_t$ sur l'intervalle de temps $[0, T_4]$, par la méthode d'Euler stochastique. (On prendra un pas de temps de $0{,}05$). On répondra à cette question en écrivant une fonction `simul_temp` qui ne prendra pas d'argument et ne renverra pas de valeur, et affichera le graphique de θ en fonction du temps, ainsi que (sur le même graphique) la courbe constante égale à Θ_{f} (cf. tableau 3).

Indication — Pour tracer deux courbes sur le même graphique, il faut écrire les commandes `plot` correspondantes entre les instructions « `hold on` » et « `hold off` ».

Le circuit fond lorsque, à un moment donné entre 0 *et* T_4, θ_t *dépasse* Θ_{f}.

5 pt **11.** Écrire une fonction `proba_fusion` estimant, par la méthode de Monte-Carlo, la probabilité que le circuit fonde, assortie d'un intervalle de confiance à 95 %. Cette fonction ne prendra aucun argument et ne renverra aucune valeur, se contentant d'afficher l'estimateur calculé et sa marge d'erreur.

10 pt **12.** Reprendre la question précédente avec un pas de temps de $0{,}01$ pour la simulation, et comparer les résultats obtenus (avec, disons, 500 000 simulations pour le pas de temps originel et 200 000 simulations pour le nouveau pas [14]). Vous devriez normalement observer que ces résultats ne sont pas compatibles... À quoi cela est-il dû, à votre avis ?

9 pt **13.** Expliquer le fonctionnement et l'intérêt du programme suivant :

```
function b = fusion_super ()
global % La fin de cette ligne est cachée.
initialiser_parameteres (); % La fonction "initialiser_parametres"
                            % initialise les paramètres globaux.
b = false;
t = 0;
temperature = temperature_initiale (); % La fonction "température initiale"
                                       % simule la loi de la température
                                       % initiale.
while t < T4
    if temperature > Tfusion
        b = true;
        break
    end
    if temperature > Tfusion - 4.
        pas = .01;
    else
        pas = .05;
    end
    temperature = temperature + % La fin de cette ligne est cachée
    t = t + pas;
end
return
end
```

On considère une nouvelle loi de probabilité $\mathbb{P}'$ *correpondant au même modèle que la probabilité* $\mathbb{P}$ *qui nous intéresse, à ceci près que* Θ_t *y est remplacée par la fonction* Θ'_t *définie*

[14]. Pour un tel nombre de simulations, les calculs peuvent être très lents ; si vous ne pouvez pas monter jusqu'à un tel niveau, expliquez alors juste comment vous effectuez votre comparaison, quitte à ce que votre conclusion s'avère différente de celle prévue par l'énoncé.

par :

$$\Theta'_t := \begin{cases} \Theta_t & \textit{pour } t \notin [T_1, T_3), \\ \Theta_t + \Delta & \textit{pour } t \in [T_1, T_3). \end{cases}$$

Nous admettons provisoirement que la loi du processus $(\theta_t)_t$ sous $\mathbb{P}'$ est alors à densité par rapport à la loi de $(\theta_t)_t$ sous $\mathbb{P}$, avec

$$\frac{d\mathbb{P}'}{d\mathbb{P}} = \exp\left(\frac{\Delta\lambda}{\sigma^2}\left(\theta_{T_3} - \theta_{T_1} + \lambda \int_{T_1}^{T_3} (\theta_t - \Theta_t - \Delta/2)\,dt\right)\right), \tag{5}$$

dont on calculera la valeur numérique par la méthode des rectangles.

9 pt **14.** Améliorer la fonction `proba_fusion` en une fonction `proba_fusion_pref` utilisant une technique d'échantillonnage préférentiel, en prenant $\mathbb{P}'$ pour loi d'échantillonnage (avec un pas de temps de $0{,}05$).

8 pt **15.** Que se serait-il passé si on avait pris $\Delta = 20$? Expliquer.

9 pt **16.** Démontrer (informellement) la formule (5).

Formulaire

Lettres grecques

Le tableau 4 donne le nom des lettres grecques utilisées dans cet énoncé.

Nom	Majuscule	Minuscule
alpha	A	α
delta	$\varDelta$	δ
epsilon	E	ε
lambda	$\varLambda$	λ
phi	$\varPhi$	φ
sigma	$\varSigma$	σ
thêta	$\varTheta$	θ

TABLEAU 4 – Quelques lettres grecques

Trigonométrie

Le *degré (sexagésimal)*, noté « ° », est une unité de mesure d'angle telle que

$$1^\circ = \frac{\pi}{180}\ \mathrm{rad}.$$

Rappelons que le radian est une unité de mesure sans dimension, càd. de même homogénéité physique que les purs nombres ; et qu'en l'absence d'unité de mesure, par défaut les angles se mesurent toujours en radians.

La fonction sinus, notée $\sin$, s'obtient sous MATLAB par la commande « `sin` » ; la fonction cosinus, notée $\cos$, s'obtient par la commande « `cos` ». On a les propriétés suivantes, pour tout $x \in \mathbf{R}$:

$$\sin(x+\pi/2) = \cos(x), \quad \cos(x+\pi/2) = -\sin(x)$$
$$\sin(x+\pi) = -\sin(x), \quad \cos(x+\pi) = -\cos(x)$$
$$\sin(x+2\pi) = \sin(x), \quad \cos(x+2\pi) = \cos(x)$$
$$\frac{d}{dx}\sin(x) = \cos(x), \quad \frac{d}{dx}\cos(x) = -\sin(x)$$
$$\sin(x)^2 + \cos(x)^2 = 1$$
$$\sin(0) = 0, \quad \cos(0) = 1$$
$$x \in (0,\pi/2) \ \Rightarrow\ \sin(x), \cos(x) > 0$$
$$\sin(-x) = -\sin(x), \quad \cos(-x) = \cos(x)$$

La fonction *arc-sinus*, notée $\arcsin$ — qui s'obtient sous MATLAB par la commande « `asin` » —, est la fonction de $[-1,1]$ dans $[-\pi/2,\pi/2]$ définie comme la bijection réciproque de la fonction sinus (restreinte à $[-\pi/2,\pi/2]$), c.-à-d. que $\arcsin(x)$ est le nombre $\theta \in [-\pi/2,\pi/2]$ tel que $\sin(\theta) = x$. De même, la fonction *arc-cosinus*, notée $\arccos$ — qui s'obtient sous MATLAB par la commande « `acos` » —, est la fonction de $[-1,1]$ dans $[0,\pi]$ définie comme la bijection réciproque de la fonction cosinus, c.-à-d. que $\arcsin(x)$ est le nombre $\theta \in [0,\pi]$ tel que $\cos(\theta) = x$. On a les propriétés suivantes :

Les fonctions $\arcsin$ et $\arccos$ sont continues.

$\arcsin$ est strictement croissante ; $\arccos$ est strictement décroissante.

$$\text{pour } x \in (-1,1), \quad \frac{d}{dx}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}, \quad \frac{d}{dx}\arccos(x) = -\frac{1}{\sqrt{1-x^2}}$$

$$\forall x \in [-1,1] \quad \arcsin(x) + \arccos(x) = \pi/2$$

$$\arcsin(0) = 0, \quad \arccos(0) = \pi/2$$

$$\arcsin(1) = \pi/2, \quad \arccos(1) = 0$$

$$\arcsin(-x) = -\arcsin(x), \quad \arccos(-x) = \pi - \arccos(x)$$

Sous MATLAB, la valeur de π peut s'obtenir par la commande « `pi` ».

Probabilités

On rappelle que la variance d'une variable aléatoire X peut être définie comme $\mathbb{E}\big((X - \mathbb{E}(X))^2\big)$.

La notation $\mathcal{N}(a,b)$ désigne la loi normale de moyenne a et de variance b. La loi normale *standard* est la loi $\mathcal{N}(0,1)$; on peut la réaliser sous MATLAB à l'aide de la commande « `randn` ». Pour cette loi, chacun des évènements suivants se produit avec une probabilité pratiquement égale à 5 % :

- Être inférieure à $-1{,}65$;
- Être supérieure à $1{,}65$;
- Être supérieure à $1{,}96$ en valeur absolue.

Un mouvement brownien est dit *standard* lorsque sa variance par unité de temps est égale à 1.

La commande « `rand` » de MATLAB simule une variable aléatoire uniforme sur $(0,1)$, indépendante à chaque appel.