TP de Modélisation Stochastique n^0 4 : Chaînes de Markov (fin) Méthode de Monte-Carlo

Ce TP se déroule en trois temps: d'abord l'installation (?) de STIXBOX, puis un exercice de chaîne de Markov (avec solution), et enfin des calculs de volumes par la méthode de Monte-Carlo.

Implantation de Stixbox

cd TP	Se placer dans ce répertoire
$cp -R \sim bardet/stixboxA111 \ \ stixboxA111$	Copier dans le répertoire TP un répertoire stixboxA111 qui comporte les outils MATLAB supplémentaires crées pour l'épreuve.
cd ..	Revenir dans le répertoire principal.
emacs .cshrc &	Ouvrir l'éditeur UNIX. Se placer dans cette fenêtre et rajouter n'importe où la ligne suivante:
$setenv\ MATLABPATH \ \sim agreg@@/stixboxA111$	(donne le chemin pour aller dans stixbox). Quitter emacs.
$cd\ stixboxA111$	Se placer dans ce nouveau répertoire.
ls	Lister les fichiers composant stixboxA111
emacs stixdemo.m &	Editer le contenu du fichier stixdemo.m
matlab	Lancer MATLAB
stixdemo	Lancer le programme stixdemo. Le faire tourner.

Un exercice de chaîne de Markov: le modèle d'Ehrenfest

On considère deux urnes contenant un nombre total n de boules que l'on choisit, avec un nombre initial k (choisi aussi) de boules dans l'urne 1. La chaîne de Markov $(X(i))$ représente le nombre de boules dans cette urne, sachant que si à l'étape i, on a $X(i) = j$, alors on enlève une boule de l'urne 1 avec la probabilité j/n et on la rajoute dans l'urne 2 (réciproquement avec une probabilité de $1-(j/n)$). Écrire un programme ($Markov-Ehrenfest.m$) permettant d'estimer la mesure invariante par le théorème ergodique (mesure invariante empirique) et on la comparer avec la mesure invariante théorique (binomiale $(n, 0.5)$). On vérifie la convergence vers la mesure invariante quelque soit le point de départ de la chaîne.

On peut consulter un exemple de ce programme situé dans stixboxA111 et executable sous le nom $Markov$.

Exercices sur la méthode d'approximation de Monte-Carlo

1. On veut trouver une valeur approchée de π. Pour cela, utiliser la méthode d'approximation de Monte-Carlo pour l'intégrale $\int_0^1 4\sqrt{1-x^2}dx$ et calculer la précision de votre algorithme (faire varier la taille de l'échantillon).
2. Trouver une approximation de la surface du disque de rayon 1 dans $\mathbb{R}^2$ (c'est-à-dire π) en utilisant la méthode de Monte-Carlo: on génére N couples de v.a.i.i.d. suivant une loi uniforme sur $[-1,1]$ (on remplit ainsi le carré de rayon 2 et de centre O) et on calcule la proportion de points dans le cercle de rayon 1 (ceci revient au calcul de $\int_{-1}^1 \int_{-1}^1 \mathbb{I}_{\{x^2+y^2\leq 1\}}(x,y)dxdy$). Faire varier N et donner un intervalle de confiance à 95% de l'estimation.
3. Reprendre l'exercice précédent dans $\mathbb{R}^3$.

4. Généralisation: On considère une fonction f continue sur un intervalle $[a, b]$ de $\mathbb{R}$, et à valeurs dans $\mathbb{R}^+$. On se propose d'évaluer numériquement l'intégrale I définie par : $I = \int_a^b f(x)dx$. A titre d'exemple, on évaluera l'intégrale J entre 0 et 2π de la fonction f définie par $f(x) = 1 + \cos(x)e^{-0.2*x}$ (on peut d'ailleurs calculer la valeur exacte de cette intégrale ...).

   a/ Méthodes déterministes 1/ On peut approcher I par une intégrale de Riemann, en discrétisant l'intervalle $[a, b]$. On peut alors arbitrairement construire : $I_1(n) = \sum_{i=1}^{n-1}(x_{i+1} - x_i)f(x_i)$ et $I_2(n) = \sum_{i=1}^{n-1}(x_{i+1} - x_i)f(x_{i+1})$, où $x_1 = a$ et $x_n = b$. En choisissant une discrétisation régulière, calculer la précision de $I_1(n)$ et $I_2(n)$ en fonction de n (supposer f dérivable).
   2/ On peut améliorer l'approximation, en approchant I par des trapèzes (méthode des trapèzes). Comparer l'approximation $I_3(n)$ obtenue avec cette méthode à $I_1(n)$ et $I_2(n)$. Calculer J avec ces trois méthodes en choisissant n de manière à avoir une précision de 10^{-3}.

   b/ Méthodes de Monte-Carlo 1/ Utiliser la méthode du premier exercice pour approcher I et déterminer n pour calculer J à 10^{-3} près avec un niveau de confiance de 95%.
   2/ On définit un rectangle R de cotés $[a, b] \times [c, d]$ tel que $c \leq f(x) \leq d$ pour tout $a \leq x \leq b$. On choisit alors aléatoirement n points indépendants, avec une distribution uniforme dans le rectangle R. Quelle est la probabilité pour qu'un de ces n points soit entre la courbe de f et l'axe des abscisses ? En déduire un estimateur $\hat{I}_n$ de I et calculer la variance de I_n (risque quadratique). Utiliser cette méthode pour évaluer J, en utilisant différentes tailles d'échantillon n. Donner à chaque fois un intervalle de confiance à 95%. Que doit valoir n pour avoir une valeur approchée à 10^{-3} avec un niveau de confiance à 95%?