Informatique

TP 8

Méthode de Monte-Carlo

Partie 1 : principe et premiers exemples

On considère une fonction $g : \mathbb{R} \longrightarrow \mathbb{R}$ et une variable aléatoire X. On suppose que $g(X)$ est une variable aléatoire admettant une espérance m et une variance σ^2.
La loi des grands nombres assure qu'étant donnée une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes et de même loi que X, la suite de terme général

$$\overline{g}_n = \frac{1}{n}\sum_{k=1}^{n} g(X_k)$$

converge en probabilité vers $m = \mathbb{E}\big(g(X)\big)$.
La méthode de Monte-Carlo consiste à approcher la quantité $\mathbb{E}\big(g(X)\big)$, difficile à calculer, par une réalisation de la variable aléatoire $\overline{g}_n$ obtenue par simulation pour n assez grand.

Exercice 1

Utiliser la méthode de Monte-Carlo pour déterminer une valeur approchée des séries et intégrales ci-dessous :

$$S = \sum_{n=1}^{\infty} \frac{\sqrt{n}}{2^n}, \qquad I = \int_{-2}^{3} \mathbf{e}^{-t^2}\,\mathrm{d}t \qquad \text{et} \qquad J = \int_{0}^{+\infty} \frac{\mathbf{e}^{-t}}{1+t^4}\,\mathrm{d}t.$$

Partie 2 : approximation d'une probabilité

La méthode de Monte-Carlo est en particulier utilisée pour approcher la probabilité d'un événement A susceptible de se produire au cours d'une expérience. On simule pour cela l'expérience un grand nombre de fois et on considère les variables aléatoires

$$X_k = \begin{cases} 1 & \text{si l'événement A se produit au cours de la } k\text{-ième expérience} \\ 0 & \text{sinon} \end{cases}$$

qui suivent toutes la loi de Bernoulli de paramètre $p = \mathbb{P}(A)$. Si g désigne la fonction identité, alors $\overline{g}_n$ est la fréquence de réalisation de l'événement A au cours des n premières expériences, qui converge en probabilité d'après la loi des grands nombres vers $\mathbb{E}(X_1) = p = \mathbb{P}(A)$ lorsque $n \to \infty$.

Exercice 2

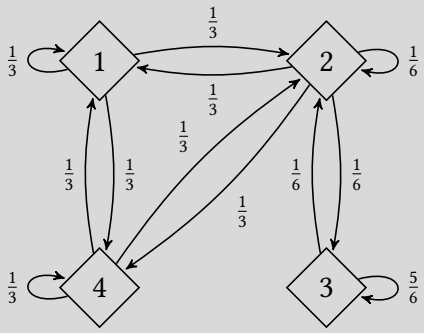
On considère le code suivant :

```
n=10^4;
X=rand(2,n);
Y=X(1,:).^2+X(2,:).^2;
disp(4*sum(Y<1)/n);
```

Que va afficher le code ci-dessus ? L'expliquer et le vérifier expérimentalement.

Exercice 3 : une chaîne de Markov

On considère les quatre sommets numérotés 1, 2, 3 et 4 d'un carré sur lesquels évolue aléatoirement une puce. Les probabilités de passage d'un sommet à l'autre sont données dans le graphe ci-dessous :



Pour tout $n \in \mathbb{N}$, on note X_n la variable aléatoire donnant la position de la puce après le n-ième saut. On considère que la puce est initialement sur le sommet 1, i.e. que $X_0 = 1$.
On définit, pour tout $n \in \mathbb{N}$, le vecteur ligne

$$L_n = \begin{pmatrix} \mathbb{P}(X_n = 1) & \mathbb{P}(X_n = 2) & \mathbb{P}(X_n = 3) & \mathbb{P}(X_n = 4) \end{pmatrix} \in \mathbf{M}_{1,4}(\mathbb{R})$$

donnant la loi de la variable aléatoire X_n.

1. Justifier que la matrice de transition

$$Q = \frac{1}{6}\begin{pmatrix} 2 & 2 & 0 & 2 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 5 & 0 \\ 2 & 2 & 0 & 2 \end{pmatrix} \in \mathbf{M}_4(\mathbb{R}),$$

dont le coefficient générique $q_{i,j}$ est la probabilité de passage du sommet i au sommet j, vérifie la relation :

$$\forall n \in \mathbb{N}, \quad L_{n+1} = L_n Q.$$

2. Déterminer deux valeurs propres « évidentes » de Q. Justifier.

3. Calculer

$$Q\begin{pmatrix} 1 \\ -1+\sqrt{2} \\ -1-\sqrt{2} \\ 1 \end{pmatrix}$$

et en déduire les deux valeurs propres restantes de Q.

4. **a.** Justifier que la suite (L_n) admet une limite Π que l'on déterminera.
b. Comment interpréter ce résultat pour la suite (X_n) ?

5. Représenter une trajectoire de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$. Observe-t-on une convergence ? Est-ce en contradiction avec le résultat de convergence établi à la question **4.** ?

6. Exécuter le code ci-dessous, identifier les courbes représentées et expliquer la convergence observée :

Listing 1 : code à analyser

```
d=50;    // durée de l'expérience
n=10^4;    // nombre de simulations
X=grand(d,'markov',Q,ones(n,1));
F=zeros(4,d);
for i=1:4
    F(i,:)=sum(X==i,'r')/n;
end
plot2d(1:d,F');
```

Exercice 4 : illustration d'un résultat démontré dans EML 2012

Pour tout entier $n \geqslant 2$, on considère une urne U_n contenant n boules numérotées de 1 à n. On effectue dans cette urne U_n des tirages successifs d'une boule avec remise. On suppose que tous les tirages dans U_n sont équiprobables. On s'arrête dès que l'on obtient une boule déjà obtenue lors d'un précédent tirage. On note T_n la variable aléatoire égale au nombre de tirages effectués.
On souhaite illustrer avec Scilab la convergence en loi de la suite de variables aléatoires de terme général $Y_n = \frac{T_n}{\sqrt{n}}$ vers une variable Y de densité

$$f : x \longmapsto \begin{cases} 0 & \text{si } x \leqslant 0 \\ x\mathrm{e}^{-x^2/2} & \text{si } x > 0 \end{cases}.$$

1. Écrire une `function y=simuleY(n)` qui simule la variable aléatoire Y_n.
2. Comment approcher la fonction de répartition de Y_n par la méthode de Monte-Carlo ? La représenter graphiquement sur $[0, 4]$ pour $n = 50$.
3. Représenter sur $[0, 4]$ la fonction de répartition de Y. Bien qu'on sache faire le calcul, on approchera la valeur de $F(x) = \int_0^x f(t)\,\mathrm{d}t$ par la méthode des rectangles pour l'exercice.
4. Illustrer la convergence en superposant les représentations précédentes pour différentes valeurs de n.

Partie 3 : garanties d'approximation

La méthode de Monte-Carlo étant basée sur des simulations, la qualité du résultat renvoyé est variable. Il n'est pas possible de faire en sorte que le résultat renvoyé soit toujours satisfaisant, mais seulement de limiter la probabilité qu'il ne le soit pas.
On dispose pour cela de deux outils :

- L'inégalité de Bienaymé-Tchebychev, qui donne

$$\mathbb{P}\left(|\overline{g}_n - m| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

- Le théorème limite central, selon lequel

$$\mathbb{P}\left(|\overline{g}_n - m| \geqslant z\frac{\sigma}{\sqrt{n}}\right) = \mathbb{P}(|\overline{g}_n^*| \geqslant z) \xrightarrow[n\to\infty]{} \mathbb{P}(|N| \geqslant z)$$

  où N suit la loi $\mathcal{N}(0, 1)$. Pour $z \simeq 1{,}96$, on obtient $\mathbb{P}(|N| \geqslant z) \simeq 5\%$.

En pratique, le théorème limite central donne de meilleurs résultats.

Exercice 5

On considère de nouveau la méthode de Monte-Carlo utilisée dans l'exercice 2 pour approcher $\frac{\pi}{4}$.

1. Quelle valeur de n choisir pour que $\overline{g}_n$ soit une valeur approchée de $\frac{\pi}{4}$ à la précision $\varepsilon = 10^{-2}$ avec un risque $\alpha = 5\%$?
2. Comment vérifier expérimentalement le résultat précédent ? Mettre cette méthode en pratique.