T.P. : Chaînes de Markov

Pour calculer explicitement la loi d'une chaîne de Markov, on se sert en général du calcul matriciel, pour lequel matlab est bien adapté. Rappel : une probabilité invariante vérifie $\mu P = \mu$: le vecteur **colonne** ${}^t\mu$ est donc un vecteur propre pour la **transposée** ${}^tP$; de plus ce vecteur doit être une probabilité, et donc vérifier $\sum_i \mu_i = 1$.

Pour simuler la chaîne, on passera a priori par l'écriture d'une fonction qui prend X_n en paramètre et simule X_{n+1} : on peut ainsi facilement simuler une trajectoire $(X_1, \dots, X_N)$. Si on a besoin d'un échantillon de X_N, il suffit de recommencer...

Exercice 1. M. Shadock peut être dans trois états : 1 (bonne santé), 2 (enrhumé), 3 (malade). Son état le jour $n+1$ dépend de son état au jour n et pas des jours précédents : s'il est en bonne santé, il le reste le lendemain avec probabilité $\frac{5}{6}$, il devient malade avec probabilités $\frac{1}{12}$. S'il est malade, il le reste avec probabilité $\frac{3}{4}$ et devient en bonne santé avec probabilité $\frac{1}{4}$. S'il est enrhumé, il guérit avec probabilité $\frac{1}{4}$ et devient malade avec probabilité $\frac{1}{4}$.

(1) Écrire la matrice de transition de la chaîne qui peut modéliser la situation. La chaîne est-elle irréductible ? Apériodique ?

(2) Déterminer sa probabilité invariante (il s'agit donc de chercher un vecteur propre colonne pour la transposée de la matrice de transition ; vous pouvez par exemple utiliser la fonction **eig**).

(3) On suppose que le premier janvier M. Shadock est en bonne santé. Quel vecteur μ_0 correspond à cette loi initiale ? Prendre un vecteur μ_0 de votre choix pour la loi initiale (la somme des probabilités doit quand même faire 1 !) Calculer la loi au n-ième jour μ_n pour $n = 5, 10, 50, 100$ et observer la convergence en loi.

(4) Recommencer avec une loi μ_0 différente. Observe-t-on encore une convergence en loi ?

(5) Compléter la fonction ci-dessous pour qu'elle permette de simuler la chaîne :

```
function Y = Shadock( X )
U=rand(1) ;
if (X==1) Y=1*(U<5/6) + 2*(U>5/6 && U<11/12) + 3*(U>11/12) ;
end
if (X==2) Y=... ; end
if (X==3) Y=... ; end
end
```

(6) Observer la convergence vers la loi invariante en générant un 5000 échantillon de la chaîne à l'instant 1000 (en partant de l'état initial que vous voulez).

Exercice 2. Zorro est en prison et dispose de 3 euros. Il peut être libéré à condition de payer une caution de 8 euros. Un gardien accepte de parier avec lui : si Zorro parie A euros, il gagne A euros avec probabilité 0.4 et perd A euros avec probabilité 0.6. Estimer numériquement la probabilité qu'il puisse payer sa caution en supposant qu'il parie 1 euro à chaque fois.

Exercice 3. Ehrenfest Soient d balles ($d > 1$) numérotées de 1 à d et réparties dans deux urnes A et B. A chaque instant, on tire un nombre i uniformément entre 1 et d, et on change la balle numéro i d'urne. Soit X_n le nombre de balles dans l'urne A après n tirages indépendants.

(1) Ecrire la matrice de transition de la chaîne en fonction de d.

(2) Vérifier à l'aide de matlab pour plusieurs valeurs de d que la loi binomiale de paramètres $(d, \frac{1}{2})$ est invariante.

(3) La chaîne est-elle irréductible ? Apériodique ?

(4) Pour $d = 5$, déterminer la loi de $X_{10}, X_{20}, X_{50}, X_{51}$ si à l'instant initial l'urne est vide. Observez-vous une convergence vers la loi invariante ?

(5) Même question (avec toujours $d = 5$) si la loi de X_0 est uniforme sur $\{0, 1, 2, 3, 4, 5\}$.

(6) On part d'une urne vide. Simuler pour $d = 5$ une trajectoire de la chaîne de Markov et tracer les $(X_1, \ldots, X_n, \ldots, X_{100})$ en fonction de n.

(7) On part d'une urne vide. Obtenir numériquement une estimation de l'espérance du nombre d'étapes nécessaire pour que l'urne soit à nouveau vide, et comparer avec sa valeur théorique (qui est $\frac{1}{P_{inv}(0)}$ où P_{inv} est la loi invariante.)

(8) Pour $d = 10$, déterminer la loi de $X_{10}, X_{20}, X_{50}, X_{51}$ si à l'instant initial l'urne est vide. A-t-on convergence vers la loi invariante ?

(9) Même question (avec toujours $d = 10$) si la loi de X_0 est uniforme.

(10) On dit que la chaîne est ergodique si pour pour tout entier m et pour tout choix de X_0, $\frac{1}{n}\sum_{k=1}^{n} 1_{X_k=m}$ converge presque sûrement quand n tend vers l'infini, et que cette limite est $P_{inv}(m)$, où P_{inv} est la loi invariante. La chaîne est-elle ergodique ? Observer numériquement le résultat.

(11) On modifie maintenant le problème : A chaque instant, on tire un nombre i uniformément entre 1 et d, et la balle numéro i change alors d'urne avec probabilité $\frac{d}{d+1}$ (et avec probabilité $\frac{1}{d+1}$, rien ne se passe).
Cette nouvelle chaîne est-elle irréductible ? Apériodique ? Reprendre les questions précédentes et comparer les résultats.

Exercice 4. Nombre de particules entrant et sortant dans un volume (X_n) désigne le nombre de particules présentes à l'instant n dans un volume V. On suppose que, pendant l'intervalle de temps $[n, n+1[$, chacune des X_n particules présentes a une probabilité $p \in]0, 1[$ de quitter V et que, pendant ce même temps, un nombre aléatoire de nouvelles particules, suivant une loi de Poisson de paramètre a, entre dans V. On suppose que les différents phénomènes aléatoires ainsi considérés sont indépendants les uns des autres.

a. Simuler une trajectoire de la chaîne pour $p = 0.1$, $a = 20$, $n = 1000$, $X_0 = 500$.
b. Représenter en fonction de k la proportion du temps entre les instants 20 et 1000 que la chaîne a passé dans chaque état k.
c. On a vu (ou on verra) que la loi de Poisson de paramètre $\frac{a}{p}$ est invariante. Comparer le graphe précédent et le diagramme en baton de cette loi de Poisson. Quelle propriété devrait être mise en évidence ?

Exercice 5.

1. Modèle de Wright-Fisher sans mutation

On modélise le nombre d'allèles A dans une population de 2N gènes par la chaîne de Markov de matrice de transition suivante, pour $0 \leq i, j \leq 2N$:

$$p_{i,j} = C_{2N}^{j} \left(\frac{i}{2N}\right)^{j} \left(1 - \frac{i}{2N}\right)^{2N-j}$$

La loi de X_{n+1} sachant que $X_n = i$ est la loi binomiale $B(2N; i/2N)$.

a) Écrire le graphe de la chaîne dans le cas $N = 2$. Quelles sont les classes irréductibles de la chaîne ? Quels sont les états absorbants ? Que signifient-ils pour la population ? La chaîne est-elle apériodique ?

b) Représentez plusieurs trajectoires de cette chaîne de Markov pour $N = 10; 20; 100$ en prenant $X_0 = N$.

c) Déterminez une estimation numérique de la probabilité que la chaîne soit absorbée en $2N$ pour $N = 10; 20; 100$ pour $X_0 = N$. Fixer $N = 10$ et tracer en fonction de k une estimation de la probabilité d'absorption en $2N$ pour $X_0 = k$. Au vu de cette courbe, proposer une formule pour cette probabilité en fonction de k.

d) Proposez une estimation de l'espérance du temps d'absorption de la chaîne pour N $= 10$; 20 ; 100, en partant de $X_0 = N$. Fixer $N = 20$ et tracer pour plusieurs valeurs de k une estimation de l'espérance du temps d'absorption pour $X_0 = k$.

e) En cherchant le sous-espace propre associé à la valeur propre 1, déterminer une mesure invariante si c'est possible.

2. Modèle de Wright-Fisher avec mutation

On modifie la chaîne en introduisant la possibilité de mutations : avant le passage à la $n+1$ème génération, chaque individu a une probabilité de muter : les individus A ont une probabilité p_{AB} de devenir B, les individus B ont une probabilité p_{BA} de devenir A. Les mutations sont supposées indépendantes les unes des autres. On obtient alors une nouvelle répartition avec Y_n individus A ; On suppose $0 < p_{AB} < 1$ et $0 < p_{BA} < 1$.

On remplace alors la chaîne précédente par le modèle suivant, avec $u = p_{AB}$ et $v = p_{BA}$ dans $]0; 1[$, pour $0 \leq i, j \leq 2N$:

$$p_{i,j} = C_{2N}^{j} \left(\frac{i}{2N}(1-u) + \left(1 - \frac{i}{2N}\right) v\right)^{j} \left(\left(1 - \frac{i}{2N}\right)(1-v) + \frac{i}{2N}u\right)^{2N-j}$$

a) Générer une trajectoire de la chaîne et la représenter. **b)** Déterminer la mesure invariante ν pour $N = 10; 20$ pour différentes valeurs de p_{AB} et p_{BA}. **c)** Proposez une méthode s'appuyant sur le théorème ergodique pour donner une estimation des coefficients de ν, i.e.

$$\frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i = l\}} \to \nu(\{l\})$$

presque sûrement. Confrontez les résultats avec ceux de la question précédentes. **d)** Obtenir ainsi pour $N = 60$ une approximation de la loi invariante π et la représenter pour différentes valeurs du couple p_{AB} et p_{BA}.

Exercice 6. Simulation d'une file d'attente à temps discret

On va modéliser une file d'attente en supposant que les événements (arrivée ou départ d'un client) ne peuvent se produire qu'à des instants entiers.
On note X_n le nombre de clients en attente ou en train d'être servis à l'instant n. Entre les instants n et $n+1$ arrivent Y_{n+1} clients, et si repartent Z_{n+1} clients (si $X_n = 0, Z_{n+1} = 0$).
On suppose que $X_0, Y_1, Z_1, Y_2, Z_2 \ldots$ sont indépendantes, que les Y_i sont indépendantes de loi de Bernoulli de paramètre p et les Z_i indépendantes de loi de Bernoulli de paramètre q.

(1) Ecrire X_{n+1} en fonction de X_n, de Y_{n+1} et de Z_{n+1}.

(2) Simuler une trajectoire de X_n et tracer les $(X_1, \ldots, X_n, \ldots X_{1000})$ en fonction de n. Prendre plusieurs valeurs de p et q et mettre en évidence des comportements différents suivant si $p < q$, $p = q$ ou $q < p$.

(3) On peut montrer que (X_n) est une chaîne de Markov irréductible apériodique. De plus si $p < q$, on admettra qu'elle admet une loi de probabilité invariante π donnée par :

$$\pi(0) = 1 - \frac{p}{q} \ \forall k \in \mathbb{N}^*, \pi(k) = (\frac{1}{1-q})(1 - \frac{p}{q})(\frac{p}{q})^k(\frac{1-q}{1-p})^k$$

On admettra que π définit bien une loi de probabilité.

(a) On dit que la chaîne est ergodique si pour pour tout entier m, $\frac{1}{n}\sum_{k=1}^{n} 1_{X_k=m}$ converge presque sûrement quand n tend vers l'infini, et que cette limite est $\pi(m)$. Tester pour différentes valeurs de p et q si la chaîne vous semble ergodique.

(b) Simuler plusieurs (au moins 2000) trajectoires de X_n pour $1 \leq n \leq 500$; tracer les fonctions de répartition empiriques de X_{10}, X_{20}, X_{50}, X_{150}, X_{500} pour différentes valeurs de p et q ainsi que et dans le cas $p < q$ comparer avec la fonction de répartition de π. Observe-t-on une convergence en loi de X_n ?