Problème du 12 Mars 2021

Durée : 3 h

L'usage des calculatrices et téléphones est interdit. Les résultats indiqués dans l'énoncé peuvent êtres utilisés pour la suite du problème. Il ne faut pas hésiter à traiter les questions dans l'ordre de son choix. On considère qu'un espace probabilisé $(\Omega, \mathcal{A}, p)$ correspond à la situation décrite.

Problème sur les urnes d'Ehrenfest.

On dispose de N boules (numérotées de 1 à N) et de deux urnes (A et B). On note $k = 0$ l'état initial des urnes : à cet état les N boules sont dans l'urne A. On s'intéresse à l'expérience aléatoire suivante. A chaque changement d'état, un expérimentateur choisit au hasard un entier entre 1 et N, attrape la boule correspondante dans l'urne où elle se situe et la change d'urne. Pour tout $k \in \mathbb{N}$ on note X_k la variable aléatoire comptant le nombre de boules dans l'urne A après k tels changements d'état. On s'intéresse à l'évolution du contenu de l'urne A et en particulier au temps de retour à l'état initial.

1. Démontrer le résultat suivant tiré du cours *Sesamaths* de Terminale.

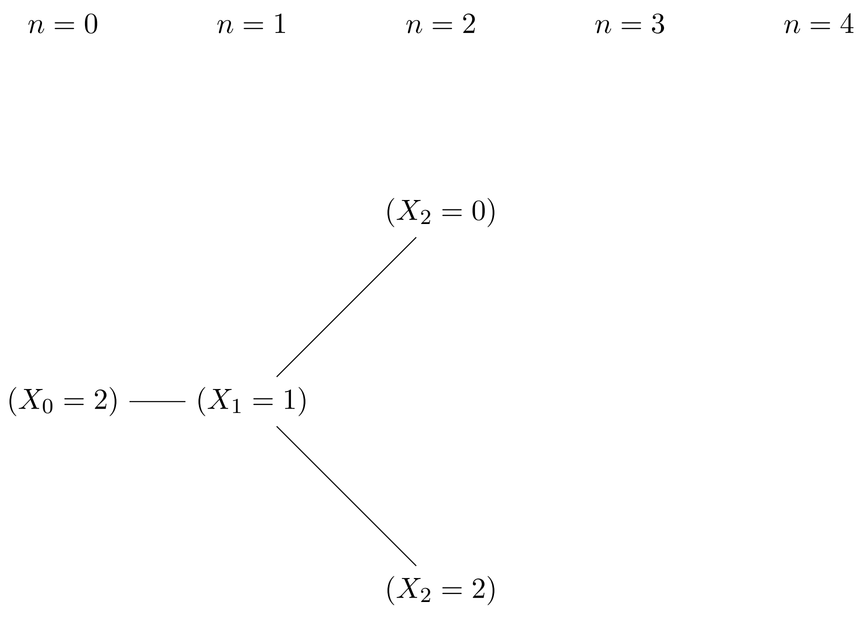
Proposition Convergence des distributions

Soit une chaîne de Markov homogène (X_n) et soit (π_n) la suite de ses distributions.
Si (π_n) est convergente alors elle converge vers une distribution invariante π_I.

2. Cas où N vaut 2.

2.1. Dans cette partie **1**, on considère que $N = 2$. Déterminer la loi de X_1, puis celle de X_2.

2.2. Compléter l'arbre de probabilités suivant jusqu'à la colonne $n = 5$.



2.3. Pour $n \in \mathbb{N}$, montrer que $(X_{2n+1} = 1) = \Omega$. (On pourra raisonner par récurrence.)

2.4. On note T la variable aléatoire donnant le nombre d'étapes nécessaires pour un premier retour à l'état initial, et prenant la valeur 0 s'il n'y a pas de retour à l'état initial. Montrer que $p\,(T = 4) = 1/4$. Puis, pour $n \in \mathbb{N}^*$, déterminer $p\,(T = 2n)$.

2.5. Calculer $\sum_{n=1}^{+\infty} p\,(T=n)$ et en déduire $p\,(T=0)$. Donner une interprétation.

2.6. Montrer que l'espérance de T est, sous réserve de convergence, égale à la somme de la série $\sum_{n\geq 0} \frac{n}{2^{n-1}}$.

2.7. Donner une justification de la convergence de cette série ne faisant pas intervenir le calcul de sa somme.

2.8. Calculer $E(T)$. (On pourra par exemple introduire une fonction $f_N(x) = \sum_{k=0}^{N} x^k$ et la dériver.)

2.9. On suppose cette fois que l'on ne connaît pas la valeur de $E(T)$, ou que l'on souhaite vérifier nos calculs. Décrire en pseudocode un algorithme simulant un grand nombre de fois l'expérience et renvoyant une valeur approchée de cette espérance.

3. Cas où N vaut 3.

3.1. Dans cette partie, on considère que $N = 3$. Représenter par un arbre probabiliste l'évolution du nombre de boules dans l'urne A au cours des quatre premières étapes, comme à la question **1.2** .

3.2. Enoncer et démontrer la formule des probabilités totales.

3.3. On considère $(U_n)_{n\in\mathbb{N}}$ la suite de matrices lignes de $\mathcal{M}_{1,4}(\mathbb{R})$ définie pour tout $n \in \mathbb{N}$ et pour tout $i \in \{0, \cdots, 3\}$ par $(U_n)_{1,i} = p\,(X_n = i)$. Déterminer sans justification une matrice $M \in \mathcal{M}_{4,4}(\mathbb{R})$ telle que $\forall n \in \mathbb{N},\ U_{n+1} = U_n M$.

3.4. Justifier votre affirmation précèdente à l'aide de la formule des probabilités totales.

3.5. Dans cette question, on suppose que la suite (U_n) est convergente. Montrer qu'il existe un seul vecteur ligne $U \in \mathcal{M}_{1,4}(\mathbb{R})$ vers lequel la suite peut converger et le déterminer.

4. Une approche algorithmique et sa justification quand $N = 3$.

4.1. Dans cette partie **3** on considère également que N vaut 3. Décrire en pseudocode un algorithme prenant en entrée un entier n non nul, un entier k, un entier i entre 0 et 3 et renvoyant en sortie une fréquence dont on espère qu'elle soit une valeur approchée de la probabilité $p\,(X_k = i)$ à l'aide de n simulations de l'expérience. Pour la suite on notera $p_{k,i}$ pour $p\,(X_k = i)$.

4.2. Montrer que la variance d'une variable aléatoire de Bernoulli est majorée par $1/4$.

4.3. On souhaite justifier l'approche utilisée à l'avant-dernière question. On notera F_n la variable aléatoire correspondant à la fréquence calculée après n simulations de l'expérience. En introduisant une suite de variables aléatoires de Bernoulli, montrer que pour tout réel $a > 0$ et tout $n \in \mathbb{N}^*$ on a

$$p\,(|F_n - p_{k,i}| \geq a) \leq \frac{1}{4a^2 n}.$$

5. Cas général.

On considère que N est un entier strictement supérieur à 3. On considère à nouveau une suite de matrices lignes à $N+1$ éléments réels $(U_n)_{n\in\mathbb{N}}$ définies pour tout $n \in \mathbb{N}$ et pour tout $i \in \{0, \cdots, N\}$ par $(U_n)_{1,i} = p\,(X_n = i)$. Déterminer en justifiant une matrice $M \in \mathcal{M}_{N+1,N+1}(\mathbb{R})$ telle que $\forall n \in \mathbb{N},\ U_{n+1} = U_n M$.