Fiche 1 – Chaînes de Markov (définition, calculs)

Quelques révisions

Exercice 1.
1. Une urne contient N boules noires et une blanche. On tire des boules successivement avec remise jusqu'à obtention de la boule blanche. On note X le nombre de tirages effectués.
1.a) Quelle est la loi de X ?
1.b) Calculer l'espérance de X.
1.c) Quelle est la probabilité qu'il faille au moins k tirages pour voir apparaître la boule blanche ?
2. Soit $n \in \mathbb{N}^*$. On suppose maintenant que l'on effectue n tirages (toujours avec remise) dans l'urne précédente, et on note B_n le nombre de fois où la boule blanche a été obtenue parmi ceux-ci.
2.a) Quelle est la loi de B_n ?
2.b) Calculer l'espérance et la variance de B_n.

Exercice 2. Soit X et Y des variables aléatoires indépendantes.
1. Quelle est la loi de $S = X + Y$ si X et Y ont pour lois respectives $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$?
2. Quelle est la loi de $Z = \min(X, Y)$ si X et Y ont pour loi respectives $\mathcal{G}(p)$ et $\mathcal{G}(q)$?

Exercice 3 – *Somme d'un nombre aléatoire de termes.* Soit X une variable aléatoire réelle intégrable. Soit $X_1, X_2, \ldots$ des variables aléatoires indépendantes, qui suivent toutes la même loi que X, et N une variable aléatoire à valeurs dans $\mathbb{N}$, indépendante des précédentes. On définit

$$S = \sum_{k=1}^{N} X_k = X_1 + \cdots + X_N.$$

1. Montrer que $\mathbb{E}[S] = \mathbb{E}[N]\mathbb{E}[X]$.
2. Donner une expression similaire pour la variance de S.

Définition des chaînes de Markov, usage en modélisation

Exercice 4 – *Lancers de dés.* On considère les jets successifs d'un dé non biaisé, autrement dit une suite $(X_n)_{n\geq 0}$ de variables aléatoires indépendantes, de loi uniforme sur $\{1, \ldots, 6\}$. On note S_n la somme et M_n le maximum des valeurs observées sur le dé entre les temps 0 et n :

$$S_n = X_0 + \cdots + X_n \qquad \text{et} \qquad M_n = \max(X_0, \ldots, X_n).$$

1. Montrer que $(X_n)_{n\geq 0}$, $(S_n)_{n\geq 0}$ et $(M_n)_{n\geq 0}$ sont des chaînes de Markov et préciser leurs probabilités de transition.
2. Que dire des suites $(S'_n)_{n\geq 0}$ et $(\Delta_n)_{n\geq 0}$ définies par

$$S'_n = X_0 + 2X_1 + \cdots + 2^n X_n \qquad \text{et} \qquad \Delta_n = X_{n+1} - X_n\,?$$

Exercice 5 – *Formalisme matriciel.* Soit $(X_n)_{n\geq 0}$ une chaîne de Markov sur $E = \{x_1, x_2, \ldots, x_N\}$, de loi initiale ν et de matrice de transition P, et $f : E \to \mathbb{R}$ bornée.
1. En termes matriciels, comment s'exprime la loi de X_1 ? de X_n ? *Assimiler les lois à des vecteurs-lignes*
2. En termes matriciels, comment s'exprime $\mathbb{E}_\nu[f(X_0)]$? $\mathbb{E}_\nu[f(X_n)]$? *Assimiler f à un vecteur-colonne*

Exercice 6 – *Météo.* À partir d'observations, un météorologue choisit de représenter le comportement journalier du temps par une chaîne de Markov à trois états : $E = \{\text{beau}, \text{nuageux}, \text{couvert}\}$, de matrice de transition

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}.$$

S'il fait beau aujourd'hui, quelle est la probabilité que le ciel soit

– couvert dans 2 jours ?
– couvert dans 2 jours et nuageux dans 4 jours ?

Exercice 7 – ***Modélisation.***
1. Dans un service administratif, on traite un dossier par jour. Par ailleurs, chaque jour arrivent zéro, un ou deux nouveaux dossiers avec probabilités p_0, p_1 et p_2 respectivement, où $p_0 + p_1 + p_2 = 1$. On note X_n le nombre de dossiers restant à traiter le jour n. Expliciter la modélisation de $(X_n)_n$ par une chaîne de Markov (Autrement dit, justifier que $(X_n)_n$ peut se représenter par une chaîne de Markov et préciser sa matrice de transition).
2. On considère un modèle (simpliste) de transmission de caractères génétiques dans un population au fil des générations. Le gène étudié possède deux allèles (« variantes ») A et a. On suppose la population de taille constante N au cours du temps et que, pour tout $n \in \mathbb{N}$, chaque individu à la génération $n+1$ hérite de l'allèle d'un des individus à la génération n choisi au hasard de façon uniforme. On note X_n le nombre d'individus possédant l'allèle A à la génération n. Expliciter la modélisation de $(X_n)_n$ par une chaîne de Markov. *Justifier que, sachant que $X_n = x$, X_{n+1} suit une loi binomiale de paramètres N et $\frac{x}{N}$*

Exercice 8 – ***Chaîne à deux états.*** Pour $p, q \in [0,1]$, on considère la chaîne sur $\{0,1\}$ de matrice de transition

$$Q = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}.$$

1. Montrer que, si $p+q > 0$, alors, pour tout $n \in \mathbb{N}$,

$$Q^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{(1-(p+q))^n}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.$$

2. En déduire $\lim_n Q^n$.
3. On suppose $\mathbb{P}(X_0 = 1) = \alpha$. Calculer $\mathbb{P}(X_0 = 1 | X_1 = 1)$.

Exercice 9 – ***Une construction des chaînes de Markov.***
1. Soit $p = (p_1, p_2, \ldots)$ un vecteur de probabilité, de taille finie ou infinie (dénombrable), c'est-à-dire que $p_i \geq 0$ pour tout i et $\sum_i p_i = 1$. On suppose que l'on dispose d'une variable aléatoire U de loi uniforme sur $[0,1]$ (par exemple fournie par un générateur de nombres pseudo-aléatoires sur un ordinateur ou une calculatrice via la fonction `rand`). Définir, à partir de U, une variable aléatoire $X = f(p, U)$ ayant la loi suivante :

$$\text{pour tout } i \in \mathbb{N}, \qquad \mathbb{P}(X = i) = p_i.$$

On pourra vérifier que $f : (p,u) \mapsto \min\{k \geq 1 | p_1 + \cdots + p_k \geq u\}$ convient.
2. Soit $E = \{x_1, x_2, \ldots\}$ un ensemble dénombrable, P une matrice stochastique sur E, et μ_0 une loi de probabilité sur E. On suppose que l'on dispose d'une suite de variables aléatoires $(U_0, U_1, U_2, \ldots)$ indépendantes et de loi uniforme sur $[0,1]$. Définir, à partir de cette suite, une suite $(X_n)_{n\geq 0}$ de variables aléatoires qui est une chaîne de Markov de mesure initiale μ_0 et de matrice de transition P.