Chaînes de Markov

Antoine Ayache

IAE de Lille

Table des matières

1 Exemples de modélisations par des chaînes de Markov

Exemple des campagnes de publicité
Une enquête réalisée sur un échantillon de consommateurs pour 3 produits P_1, P_2 et P_3, qui sont identiques mais fabriqués par 3 entreprises différentes ξ_1, ξ_2, et ξ_3, a permis de constater que :

- 30% des personnes concernées consomment le produit P_1 ;
- 50% des personnes concernées consomment le produit P_2 ;
- 20% des personnes concernées consomment le produit P_3.

Le fabricant de P_1 souhaite accroître sa part de marché ; il lance des campagnes de publicités successives. Chacune d'elles produit les effets suivants :

- parmi les clients de P_1 : 50% continuent d'acheter P_1, 40% achètent P_2, et 10% achètent P_3 ;
- parmi les clients de P_2 : 70% continuent d'acheter P_2 et 30% achètent P_1 ;
- parmi les clients de P_3 : 80% continuent d'acheter P_3 et 20% achètent P_1 ;

1) Quel sera l'état du marché au bout d'une campagne de publicité, et quel sera son état au bout de deux campagnes de publicité ?
2) A l'issue d'un certain nombre de campagnes publicitaires l'état du marché se

rapproche d'un état d'équilibre. Comment peut-on déterminer cet état d'équilibre ?
3) La vente d'un article P_1 rapporte à ξ_1 un gain de 1,5 Euros et coût pour ξ_1 d'une campagne publicitaire est de 230000 Euros. Combien de campagnes publicitaires faudrait-il que ξ_1 fasse pour obtenir un revenu maximal ?

Exemple des deux urnes

Considérons deux urnes numérotées 1 et 2. L'urne 1 contient 10 jetons sur chacun desquels figure un numéro : 6 jetons portent le numéro 1 et les 4 autres le numéro 2. L'urne 2 contient 20 jetons sur chacun desquels figure un numéro : 10 jetons portent le numéro 1 et les 10 autres le numéro 2. A l'étape 0, on tire un jeton de l'urne 1 et on lit le numéro, désigné de façon générale par i, de ce jeton, puis on replace ce dernier dans son urne d'origine (c'est-à-dire l'urne 1). A l'étape 1, on va à l'urne numéro i (parce qu'on a tiré un jeton qui porte le numéro i à l'étape 0), on y tire un jeton on lit son numéro, désigné de façon générale par j, puis on replace ce jeton dans son urne d'origine. On continue selon ce même principe dans toutes les étapes ultérieures. Ainsi, pour tout entier $t \geq 1$, lorsqu'on a tiré le jeton numéro i_{t-1} à l'étape $t-1$, à l'étape t on va à l'urne i_{t-1} et on y tire un jeton dont le numéro est désigné par i_t en toute généralité. Pour tout $t \in \mathbb{N}$, le numéro du jeton tiré à l'étape t est une variable aléatoire que nous notons par X_t ; cette variable aléatoire est à valeurs dans l'espace d'états $\mathcal{E} := \{1, 2\}$.
1) a) Calculer la probabilité p_{12} de tirer 2 à l'étape 1 lorsqu'on a tiré 1 à l'étape 0. Puis, calculer la probabilité p_{22} de tirer 2 à l'étape 1 lorsqu'on a tiré 2 à l'étape 0.
b) Calculer la probabilité p_{21} de tirer 1 à l'étape 1 lorsqu'on a tiré 2 à l'étape 0. Puis, calculer la probabilité p_{11} de tirer 1 à l'étape 1 lorsqu'on a tiré 1 à l'étape 0.
2) Calculer la probabilité de tirer 2 à l'étape 1 puis la probabilité de tirer 1 à l'étape 1.
3) On admet que, pour tout $t \geq 1$ et pour tous $i \in \mathcal{E}$, $i_{t-1} \in \mathcal{E}, \ldots, i_0 \in \mathcal{E}$, la probabilité $P(X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0)$ est non nulle.
a) Montrer que, pour tout $t \geq 1$, l'on a

$$\begin{aligned} \mathbb{P}\Big(X_{t+1} = j \big/ X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0\Big) &= \mathbb{P}\Big(X_{t+1} = j \big/ X_t = i\Big) \\ &= \mathbb{P}\Big(X_1 = j \big/ X_0 = i\Big) ; \end{aligned}$$

on commencera d'abord par étudier les cas $t = 0$, $t = 1$ et $t = 2$.
b) Que peut-on dire de la suite de variables aléatoires $(X_t)_{t \in \mathbb{N}}$?
4) Calculer la probabilité de tirer 2 à l'étape 2 puis la probabilité de tirer 1 à l'étape 2.
5) Pour tout $t \in \mathbb{N}$, on note par π_t la matrice ligne $\pi_t := \Big(\mathbb{P}(X_t = 1) \;\; \mathbb{P}(X_t = 2)\Big)$. La matrice des probabilités de transition, qu'on appelle aussi matrice de transition, est par définition la matrice $M := \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$. Montrer que l'on a, pour tout

$t \in \mathbb{N}^*$, $\pi_t = \pi_{t-1}M$, puis montrer que l'on a, pour tout $t \in \mathbb{N}$, $\pi_t = \pi_0 M^t$.
6) On admet que lorsque $p_{11} - p_{21} - 1 \neq 0$, la matrice M peut alors s'écrire sous la forme $M = QDQ^{-1}$, où la matrice $Q := \begin{pmatrix} 1 & 1 - p_{11} \\ 1 & -p_{21} \end{pmatrix}$ et Q^{-1} est son inverse, et où la matrice diagonale $D := \begin{pmatrix} 1 & 0 \\ 0 & p_{11} - p_{21} \end{pmatrix}$.

a) Après avoir justifié son existence, calculer la matrice Q^{-1} en fonction de p_{11} et de p_{21}.

b) Montrer que l'on a, pour tout $t \in \mathbb{N}$, $M^t = QD^tQ^{-1}$, puis calculer la matrice M^t en fonction de p_{11} et de p_{21}.

c) En utilisant les valeurs numériques de p_{11} et de p_{21} que vous avez trouvées en répondant à la question 1), calculer π_t, pour tout $t \in \mathbb{N}$.

d) On note par π^* la limite de π_t lorsque t tend vers $+\infty$. Calculer π^*.

e) Calculer le produit matriciel $\pi^* M$ puis interpréter votre résultat.

2 La Valeur Actualisée du Client "Customer Life Time Value"

Notre présentation du concept de la Valeur Actualisée du Client (VAC) s'inspire de celle de l'article de Calciu et Salerno [3]. Ce concept a été développé initialement par les spécialistes de la vente à distance. C'est en effet dans ce secteur que sont apparues les premières bases de données clientèle permettant le calcul de la VAC. L'analyse de la VAC permet à une entreprise de gérer son portefeuille de clientèle. Elle l'aide notamment dans la budgétisation de ses dépenses pour l'acquisition des clients, la sélection des médias de recrutement (la VAC est généralement différente selon les différents médias) ou des types d'offres, ou encore la répartition des efforts entre prospection et fidélisation. L'article [4] de Courtheaux a illustré l'utilité de la VAC dans de nombreux problèmes managériaux.

Définition 2.1 (**Valeur Actualisée du Client "Customer Life Time Value").**
Considérons un consommateur "C" qui est susceptible, au cours de différentes périodes (disons par exemple que la durée de chaque période est de trois semaines) d'acheter une unité d'un certain produit "P" (par exemple s'abonner à une certaine revue) à une entreprise "E". On suppose que "C" est présent dans la base de données de "E" au début de la période 0 *(c'est-à-dire qu'il a été préalablement à la période* 0 *un client de "E"). Désignons par :*

— $\mathcal{R}_t$ *le coût de l'effort marketing par client présent dans la base de données de "E" au début de la période t, qui est supporté par cette entreprise "E"* *lors de la période t (il ne faut pas confondre $\mathcal{R}_t$ avec le coût de production du produit "P");*

- $\mathcal{M}_t$ *le prix unitaire du produit "P" à la période t ;*
- p_t *la probabilité que "C" achète à "E" une unité de "P" à la période t ;*
- p_t^* *la probabilité que "C" soit dans la base de données de "E" au début de la période t (cette probabilité p_t^* dépend du fait que "C" a été ou n'a pas été client de "E" au cours des périodes précédentes) ;*
- a *le taux d'actualisation (on suppose qu'il reste constant, c'est-à-dire qu'il ne change pas d'une période à l'autre).*

Ayant introduit ces notations, V_T, la Valeur Actualisée du Client "C" à la période T, peut, en toute généralité, être définie par :

$$\begin{aligned} V_T &:= \sum_{t=0}^{T} \frac{\mathcal{M}_t p_t - \mathcal{R}_t p_t^*}{(1+a)^t} \qquad (2.1) \\ &= \mathcal{M}_0 p_0 - \mathcal{R}_0 p_0^* + \frac{\mathcal{M}_1 p_1 - \mathcal{R}_1 p_1^*}{(1+a)} + \ldots + \frac{\mathcal{M}_T p_T - \mathcal{R}_T p_T^*}{(1+a)^T} . \end{aligned}$$

Exemple 2.1. *1) On suppose que le consommateur "C" sera certainement présent dans la base de données de l'entreprise "E" aux débuts des périodes 0 et 1, et que "C" achètera une unité du produit "P" au cours de chacune de ces deux périodes. On suppose de plus que $\mathcal{M}_0 = 1$ Euro, $\mathcal{M}_1 = 1,2$ Euros, $\mathcal{R}_0 = 0,5$ Euro, $\mathcal{R}_1 = 0,55$ Euro, et $a = 10\%$. Calculer V_1 la VAC de "C" à la période 1.*

2) On suppose que "C" sera certainement présent dans la base de données de "E" au début de la période 0, et qu'il y a une chance sur deux qu'il achète une unité de "P" à la période 0. On suppose aussi que "C" a une chance sur deux d'être présent dans la base de données de "E" au début de la période 1, et qu'il a une chance sur quatre d'acheter une unité de "P" à la période 1. On suppose enfin que $\mathcal{M}_0 = 1,3$ Euros, $\mathcal{M}_1 = 1,5$ Euros, $\mathcal{R}_0 = 0,5$ Euro, $\mathcal{R}_1 = 0,7$ Euro, et $a = 10\%$. Calculer V_1 la VAC de "C" à la période 1.

Remarques 2.1.

- *La définition (2.1) de la Valeur Actualisée du Client se fonde sur un principe qui a de l'analogie avec celui qui permet d'évaluer une entreprise en se basant sur les bénéfices futurs actualisés qu'elle est susceptible de générer.*
- *Dans la définition (2.1) de la Valeur Actualisée du Client, à chaque période t, on sépare complètement ce que "C" devrait dépenser pour acheter une unité de "P" (c'est-à-dire $\mathcal{M}_t$) de la probabilité que "C" achète effectivement une unité de "P" (c'est-à-dire p_t), autrement dit p_t ne dépend pas de $\mathcal{M}_t$.*
- *La principale difficulté posée par le calcul de la Valeur Actualisée du Client est l'estimation, pour chaque période t, de p_t^* (la probabilité que le consommateur "C" soit dans la base de données de l'entreprise "E") et de p_t (la*

probabilité qu'il fasse un achat d'une unité du produit "P"). Le modèle de rétention et le modèle de migration sont les deux principaux modèles permettant l'estimation de ces probabilités.

Le modèle de rétention

Ce modèle considère que lorsqu'un client ne génère pas de transaction au cours d'une période, alors il ne va plus générer de transactions dans les périodes ultérieures, il est perdu pour de bon "lost for good" (pour reprendre la même terminologie que celle des articles de Dwyer [5] et de Jackson [7]), et il faut donc qu'il soit rayé de la base de données de l'entreprise dès la fin de la première période où il n'a plus généré de transaction. Ainsi, dans le cadre du modèle de rétention on a, pour tout $t \geq 1$, $p_t^* = p_{t-1}$, et en général on suppose que $p_0^* = 1$. Signalons aussi que dans le cadre du modèle de rétention, reprenant la même terminologie que celle de l'article [2] de Blattberg, Gary et Jacquelyn, pour tout $t \geq 0$, la probabilité p_t est appelée "la probabilité de survie du client jusqu'à la fin de la période t". Signalons enfin que dans le cadre du modèle de rétention, la suite de probabilités $(p_t)_{t \geq 0}$ est décroissante (pour tout $t \geq 0$, on a $p_{t+1} \leq p_t$) ; en effet, plus le temps passe moins le client a de chance de "survivre" (c'est-à-dire de générer des transactions).

Le modèle de migration

Ce modèle considère que lorsqu'un client ne génère pas de transaction au cours d'une ou même de quelques périodes, il pourrait toutefois générer des transactions ultérieurement, et il convient donc de le garder pendant un certain temps dans la base de données de l'entreprise. Les chaînes de Markov permettent de modéliser le comportement d'un client dans le cadre d'un modèle de migration (voir par exemple à ce sujet l'article de Pfeifer et Carraway [9]) ; signalons au passage qu'un tel comportement du client est du type "always a share" (pour reprendre la même terminologie que celle des articles de Dwyer [5] et de Jackson [7]).

3 Importants résultats sur les chaînes de Markov

On s'intéresse à la modélisation de l'évolution au cours du temps d'un système (par exemple l'attitude d'un client à l'égard d'un produit). On note par $\mathcal{E}$ l'ensemble de tous les états dans lesquels le système peut être. Dans le cadre de ce cours $\mathcal{E}$ est toujours un ensemble fini, et pour simplifier la présentation on suppose que $\mathcal{E} = \{1, 2, \ldots, n\}$, où n est le nombre des états. De plus, on suppose toujours que le temps est discret ; il est représenté par l'ensemble des entiers naturels $\mathbb{N} := \{0, 1, 2, 3, \ldots\}$. L'état du système en un instant $t \in \mathbb{N}$ est donné par une variable aléatoire discrète, désignée généralement par X_t, qui est à valeurs dans $\mathcal{E}$ (l'ensemble de tous les états). Intuitivement parlant, la suite de variables aléatoires $(X_t)_{t\in\mathbb{N}}$ est une chaîne de Markov, lorsqu'en chaque instant $t \geq 1$, l'évolution future du système (c'est-à-dire $X_{t+1}, X_{t+2}, X_{t+3}, \ldots$) ne dépend pas du passé (c'est-à-dire de $X_{t-1}, \ldots, X_0$) mais uniquement du présent (c'est-à-dire de X_t). De manière formelle les chaînes de Markov sont définies de la façon suivantes :

Définition 3.1 (chaîne de Markov). *Soit $(X_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires discrètes à valeurs dans un espace $\mathcal{E} = \{1, 2, \ldots, n\}$ appelé espace des états.*

1) On dit que $(X_t)_{t\in\mathbb{N}}$ est une chaîne de Markov si et seulement si

$$\mathbb{P}\Big(X_{t+1} = j \big/ X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0\Big) = \mathbb{P}\Big(X_{t+1} = j \big/ X_t = i\Big), \tag{3.1}$$

pour tout $t \geq 1$ et pour tous $i \in \mathcal{E}$, $i_{t-1} \in \mathcal{E}, \ldots, i_0 \in \mathcal{E}$ pour lesquels la probabilité $P(X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0)$ est non nulle.

2) Si en plus, la probabilité $\mathbb{P}\Big(X_{t+1} = j \big/ X_t = i\Big)$ ne dépend pas de t, c'est-à-dire si pour tout $t \in \mathbb{N}$,

$$\mathbb{P}\Big(X_{t+1} = j \big/ X_t = i\Big) = \mathbb{P}\Big(X_1 = j \big/ X_0 = i\Big), \tag{3.2}$$

on dit que la chaîne de Markov est homogène.

Remarque 3.1. *Toutes les chaînes de Markov dont il est question dans tout le reste de cette Section 3 sont supposées être homogènes bien que, la plupart du temps, cela n'est pas mentionné explicitement.*

Définition 3.2 (matrice des probabilités de transition, encore appelée matrice de transition). *C'est la matrice carrée à n lignes et à n colonnes (rappelons que n est le nombre des états) dont les coefficients sont les probabilités de transition d'un état vers un autre. Généralement, cette matrice est notée par M, et pour tous $i \in \mathcal{E}$ et $j \in \mathcal{E}$, p_{ij} désigne le coefficient qui se trouve à la i-ème ligne et la j-ème colonne de M. Il est important de retenir que*

$$p_{ij} := \mathbb{P}\Big(X_1 = j \big/ X_0 = i\Big), \tag{3.3}$$

c'est-à-dire que p_{ij} est la probabilité de passer à l'état j à la période 1 (ou à l'étape 1) lorsqu'on se trouve à l'état i à la période 0 (ou à l'étape 0).

Remarque 3.2. *La matrice de transition M est une matrice stochastique, c'est-à-dire que tous ses coefficients sont compris entre 0 et 1, et que la somme des coefficients de chaque ligne de M vaut 1.*

Remarque 3.3. *Pour tout $t \in \mathbb{N}$, nous notons par π_t la matrice (ou vecteur) ligne des probabilités d'apparition de chacun des états d'une chaîne de Markov $(X_t)_{t\in\mathbb{N}}$, à la période t, c'est-à-dire que*

$$\pi_t := \Big(\mathbb{P}(X_t = 1), \mathbb{P}(X_t = 2), \ldots, \mathbb{P}(X_t = n)\Big). \tag{3.4}$$

Signalons au passage que la loi de probabilité de la variable aléatoire X_t est complètement déterminée par le vecteur π_t.

Théorème 3.1 (ce théorème fondamental est dû à Chapman et Kolmogorov). *Soient $m \in \mathbb{N}$ et $t \in \mathbb{N}$ deux entiers naturels quelconques, pour tout $j \in \mathcal{E}$ et tout $i \in \mathcal{E}$, la probabilité d'atteindre l'état j à la période $m+t$ lorsqu'on se trouve à l'état i à la période m, autrement dit la probabilité conditionnelle $\mathbb{P}\Big(X_{m+t} = j \Big/ X_m = i\Big)$, c'est tout simplement le coefficient de la matrice M^t (M à la puissance t) situé à sa i-ème ligne et à sa j-ème colonne. En fait, à cause de l'homogénéité de la chaîne de Markov, cette probabilité conditionnelle ne dépend pas de m ; ainsi, elle est notée par $p_{ij}^{(t)}$.*

Retenir que :

— p_{ij}, le coefficient situé à la i-ème ligne et à la j-ème colonne de la matrice de transition M, est la probabilité d'atteindre l'état j en 1 seule étape, lorsqu'on se trouve à l'état i.

— Pour tout $t \in \mathbb{N}$, $p_{ij}^{(t)}$, le coefficient situé à la i-ème ligne et à la j-ème colonne de la matrice M^t, est la probabilité d'atteindre l'état j en t étape(s) lorsqu'on se trouve à l'état i. On a donc $p_{ij}^{(1)} = p_{ij}$, $p_{ii}^{(0)} = 1$ et $p_{ij}^{(0)} = 0$ lorsque $i \neq j$. Les probabilités $p_{ij}^{(t)}$ sont appelées les probabilités de transition en t étapes.

Remarque 3.4. *Pour tout $t \in \mathbb{N}$, la matrice de transition M^t est une matrice stochastique, c'est-à-dire que tous ses coefficients sont compris entre* 0 *et* 1, *et que la somme des coefficients de chaque ligne de M^t vaut* 1.

Corollaire 3.1. *1) Pour toute période $t \geq 1$, π_t le vecteur (ou matrice) ligne des probabilités d'apparition de chacun des états peut être obtenu à partir de π_{t-1} au moyen de la relation de récurrence matricielle*

$$\pi_t = \pi_{t-1} M. \tag{3.5}$$

2) Pour toute période $t \geq 0$, π_t peut être obtenu à partir de π_0 au moyen de la relation matricielle

$$\pi_t = \pi_0 M^t, \tag{3.6}$$

où la matrice M^t (M à la puissance t) est obtenue en faisant le produit de la matrice M avec elle-même t fois.

Remarque 3.5. *1) L'égalité matricielle (3.5) signifie que, pour tout $t \geq 1$ et tout $j \in \mathcal{E}$, on a*

$$\mathbb{P}(X_t = j) = \sum_{i=1}^{n} \mathbb{P}(X_{t-1} = i) p_{ij}, \tag{3.7}$$

où les p_{ij} sont les probabilités de transition en 1 *seule étape.*

2) L'égalité matricielle (3.6) signifie que, pour tout $t \geq 1$ et tout $j \in \mathcal{E}$, on a

$$\mathbb{P}(X_t = j) = \sum_{i=1}^{n} \mathbb{P}(X_0 = i) p_{ij}^{(t)}, \tag{3.8}$$

où les $p_{ij}^{(t)}$ sont les probabilités de transition en t étapes.

Définition 3.3 (graphe d'une chaîne de Markov). *Le graphe d'une chaîne de Markov est obtenu en traçant, pour tout $i \in \mathcal{E}$ et tout $j \in \mathcal{E}$, une flèche allant de i vers j, lorsque p_{ij} la probabilité de transition de l'état i vers l'état j est* **non nulle** *; on met généralement à côté de cette flèche la valeur de p_{ij}.*

Le graphe d'une chaîne de Markov permet de détecter les états qui sont accessibles l'un à partir de l'autre, et ceux qui communiquent entre eux.

Définition 3.4 (accessibilité d'un état à partir d'un état, et communication de deux états). *On dit qu'un état j d'une chaîne de Markov est accessible à partir d'un état i de celle-ci (notons au passage que l'état j pourrait éventuellement être l'état i lui-même), lorsque dans le graphe de la chaîne de Markov il existe au moins un chemin qui permet d'aller de i vers j (ce chemin peut bien sûr comporter plusieurs flèches et pas seulement une seule).*

On dit que deux états i et j d'une chaîne de Markov communiquent entre eux (notons au passage que l'état j pourrait éventuellement être l'état i lui-même), lorsque j est accessible à partir de i, et que inversement l'on a aussi que i est accessible à partir de j.

Définition 3.5 (état absorbant). *Un état i d'une chaîne de Markov est dit absorbant, lorsque aucun autre état $j \neq i$ de la chaîne de Markov n'est accessible à partir de i.*

Définition 3.6 (chaîne de Markov irréductible). *Une chaîne de Markov est dite irréductible lorsque tous ses états communiquent les uns avec les autres.*

Définition 3.7 (Probabilité invariante). *Soit une matrice (ou vecteur) ligne $\pi^* = (p_1^*, p_2^*, \ldots, p_n^*)$ telle que $p_1^*, p_2^*, \ldots, p_n^*$ sont n nombres réels tous compris entre 0 et 1, et dont la somme vaut 1 (c'est-à-dire $\sum_{j=1}^{n} p_j^* = 1$). On dit que π^* est une probabilité invariante d'une chaîne de Markov à n états et de matrice de transition M lorsque l'on a*

$$\pi^* = \pi^* M. \tag{3.9}$$

La proposition suivante montre que π^* correspond à une sorte d'équilibre.

Proposition 3.1. *Rappelons que (voir la Remarque 3.3) que, pour tout $t \in \mathbb{N}$, π_t désigne la matrice (ou vecteur) ligne des probabilités d'apparition de chacun des* *états d'une chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ à la période t. Supposons qu'à une certaine période t_0, on a $\pi_{t_0} = \pi^*$, où π^* est une probabilité invariante de $(X_t)_{t\in\mathbb{N}}$. Alors à toutes les périodes t qui sont ultérieures à t_0 (c'est-à-dire pour tout $t \geq t_0$), on a nécessairement $\pi_t = \pi^*$.*

Théorème 3.2 (**trois propriétés "sympathiques" des chaînes de Markov irréductibles**). *Soient $(X_t)_{t\in\mathbb{N}}$ une chaîne de Markov* **irréductible**, *$M = (p_{ij})_{1\leq i,j\leq n}$ sa matrice de transition et $\mathcal{E} = \{1, 2, \ldots, n\}$ son espace d'états.*

1) La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ admet **une unique probabilité invariante** *$\pi^* = (p_1^*, p_2^*, \ldots, p_n^*)$ (c'est-à-dire que $\pi^* = \pi^* M$), de plus $p_1^*, p_2^*, \ldots, p_n^*$ sont tous strictement compris entre 0 et 1, ce qui notamment signifie qu'ils sont tous non nuls.*

2) Quelque soit l'état initial $i \in \mathcal{E}$, la fréquence d'apparition moyenne d'un état arbitraire $j \in \mathcal{E}$ converge vers p_j^*. Plus précisément, pour tout $i \in \mathcal{E}$ et tout $j \in \mathcal{E}$,

$$p_j^* = \lim_{T\to+\infty} \frac{1}{T} \sum_{t=1}^{T} p_{ij}^{(t)},$$

où, pour tout $t \geq 1$, $p_{ij}^{(t)}$ désigne la probabilité de transition en t étape(s) de l'état i vers l'état j.

3) Supposons qu'il existe $l \in \mathcal{E}$ tel que $p_{ll} > 0$ (c'est-à-dire que la probabilité de rester dans un certain état l d'une étape à une autre est non nulle). Alors, pour tout $i \in \mathcal{E}$ et tout $j \in \mathcal{E}$,

$$p_j^* = \lim_{t\to+\infty} p_{ij}^{(t)};$$

ce qui signifie que quelque soit l'état initial i, $p_{ij}^{(t)}$, la probabilité de transition en t étape(s) de l'état i vers un état arbitraire j, converge vers p_j^* lorsque t tend vers $+\infty$.

Remarque 3.6. *Il résulte de la Partie 2) du Théorème 3.2 et de la formule des probabilités totales, que la fréquence moyenne d'apparition d'un état arbitraire $j \in \mathcal{E}$ converge vers p_j^*. Plus précisément, on a*

$$p_j^* = \lim_{T\to+\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}(X_t = j);$$

ce qui signifie qu'on peut interpréter p_j^* comme la fréquence moyenne d'apparition de l'état j à long terme.

Remarque 3.7. *Il résulte de la Partie 3) du Théorème 3.2 et de la formule des probabilités totales,que*

$$p_j^* = \lim_{t\to+\infty} \mathbb{P}(X_t = j);$$

ce qui signifie qu'on peut également interpréter p_j^* comme la probabilité de l'état j à long terme.

4 Exercices

Exercice 1
Le chauffage d'une maison individuelle est composé d'un chauffage de base et d'un chauffage d'appoint. On dit qu'on est dans l'état 1 si seul le chauffage de base fonctionne, et on dit qu'on est dans l'état 2 si les deux chauffages sont en marche. Chaque jour on se trouve dans l'un de ces deux états et on ne peut changer d'état que le lendemain. L'état où l'on se trouve à la date t est donné par la variable aléatoire X_t qui est à valeurs dans l'espace d'états $\mathcal{E} := \{1, 2\}$. On admet que la suite $(X_t)_{t\in\mathbb{N}}$ est une chaîne de Markov homogène dont les probabilités de transition d'un état vers un autre sont déterminées de la manière suivante : si on est dans l'état 1 un jour, on estime que la probabilité de rester dans l'état 1 le lendemain vaut $1/2$; si on contraire, on est dans l'état 2 un jour, on estime que le lendemain la maison sera encore bien chaude et ainsi qu'on passera à l'état 1 avec probabilité $3/4$.
1) Donner la matrice de transition M de la chaîne de Markov homogène $(X_t)_{t\in\mathbb{N}}$ puis tracer son graphe.
2) Déterminer les états qui sont accessibles l'un à partir de l'autre et ceux qui communiquent entre eux. Y-a-t-il un état absorbant. La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ est-elle irréductible ?
3) Pour tout $t \in \mathbb{N}$, on note par π_t la matrice ligne $\pi_t := \Big(\mathbb{P}(X_t = 1)\ \ \mathbb{P}(X_t = 2)\Big)$; on admet que $\pi_0 = (1\ \ 0)$, ce qui signifie concrètement qu'on se trouve nécessairement à l'état 1 à la date initiale $t = 0$.

a) Calculer π_1 et π_2.

b) Plus généralement, calculer π_t, pour tout $t \in \mathbb{N}$.

4) Montrer que La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ possède une unique probabilité invariante π^*, puis calculer π^*.
5) Rappelons qu'à la date initiale $t = 0$ on se trouve dans l'état 1. La variable aléatoire τ désigne la date à laquelle on passe pour la première fois à l'état 2.

a) Calculer les probabilités $\mathbb{P}(\tau = 1)$, $\mathbb{P}(\tau = 2)$ et $\mathbb{P}(\tau = 3)$.

b) Plus généralement, pour tout $n \in \mathbb{N}^*$, calculer la probabilité $\mathbb{P}(\tau = n)$.

c) Trouver la loi de probabilité de la variable aléatoire τ puis calculer son espérance, sa variance et son écart-type.

Exercice 2
Les trois questions de cet exercice sont indépendantes l'une de l'autre.
1) On désigne par $(X_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires discrètes à valeurs dans l'ensemble $\{1, 2, 3, 4\}$. On sait que les probabilités conditionnelles $\mathbb{P}(X_4 = 3/X_1 = 1, X_3 = 2)$ et $\mathbb{P}(X_4 = 3/X_3 = 2)$ valent respectivement $1/4$ et $1/2$; est-ce que la suite $(X_t)_{t\in\mathbb{N}}$ peut être une chaîne de Markov ? (justifier votre réponse)
2) On désigne par $(X_t)_{t\in\mathbb{N}}$ une chaîne de Markov homogène à valeurs dans l'espace

d'états $\{1,2,3\}$. On suppose que les deux probabilités conditionnelles $\mathbb{P}(X_1 = 1/X_0 = 2)$ et $\mathbb{P}(X_1 = 3/X_0 = 2)$ valent zéro toutes les deux. Que peut-on en conclure ? (justifier votre réponse)
3) On sait que M est une matrice carrée d'ordre 3, vérifiant :

$$M^2 = \begin{pmatrix} 1/9 & 1/3 & 1/4 \\ 0 & 1/9 & 1/3 \\ 0 & 0 & 1/9 \end{pmatrix}.$$

M peut-elle être la matrice de transition d'une chaîne de Markov homogène à trois états ? (justifier votre réponse)

Exercice 3
Dans cet exercice le temps est divisé en périodes. Considérons un consommateur "C" qui est susceptible d'acheter à chaque période une unité d'un certain produit "P" à une certaine entreprise "ξ".
1) On suppose que le comportement de ce consommateur peut être modélisé par une chaîne de Markov homogène $(X_t)_{t\in\mathbb{N}}$ d'espace d'états $\mathcal{E} := \{1,2,3\}$ et dont la matrice de transition est

$$M = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 0 & 3/4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Le consommateur "C" fera un achat à une période t lorsque $X_t = 1$ et sinon il ne fera pas d'achat.
a) Tracer le graphe de la chaîne de Markov $(X_t)_{t\in\mathbb{N}}$.
b) Déterminer les états qui sont accessibles l'un à partir de l'autre et ceux qui communiquent entre eux. Y-a-t-il un état absorbant. La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ est-elle irréductible ?
2) A la fin de chaque période, le service marketing de "ξ" décide de garder "C" dans la base de données tant que celui-ci est encore susceptible d'acheter "P" dans des périodes ultérieures. Dans le cas contraire "C" est définitivement rayé de la base de données. Pour tout $t \in \mathbb{N}$ on désigne par p_t^* la probabilité que "C" soit dans la base de données de "ξ" au début de la période t. On suppose que $p_0^* = 1$. Pour tout $t \in \mathbb{N}$ on désigne par p_t la probabilité que "C" achète une unité de "P" à la période t. On suppose que $p_0 = 1$.
a) Montrer pour tout $t \geq 1$ on a $p_t^* = 1 - \mathbb{P}(X_{t-1} = 3)$.
b) Calculer p_1, p_2 et p_3.
c) Calculer p_1^*, p_2^* et p_3^*.
3) La variable aléatoire τ désigne le nombre aléatoire d'unité(s) de "P" qui ont été achetée(s) par "C" au cours de toutes les périodes. Calculer les probabilités

$\mathbb{P}(\tau = 1)$ et $\mathbb{P}(\tau = 2)$.
4) Pour tout $t \geq 0$, on désigne par $\mathcal{R}_t$ le coût de l'effort marketing par client présent dans la base de données de "ξ" au début de période t et on désigne par $\mathcal{M}_t$ le prix unitaire du produit "P" à la période t. Les quantités $\mathcal{M}_t$ et $\mathcal{R}_t$ sont exprimés en Euro(s) de la période t. On suppose que le taux d'actualisation reste constant d'une période à une autre. Ce taux est noté par r.

a) Exprimer V_3 la valeur actualisée du client "C" à la période 3 au moyen de $p_0, \ldots, p_3$, $p_0^*, \ldots, p_3^*$, $\mathcal{M}_0, \ldots, \mathcal{M}_3$, $\mathcal{R}_0, \ldots, \mathcal{R}_3$ et r.

b) Calculer V_3 lorsque $\mathcal{R}_0 = \mathcal{R}_1 = \mathcal{R}_2 = \mathcal{R}_3 = 1$ Euro, $\mathcal{M}_0 = \mathcal{M}_1 = \mathcal{M}_2 = \mathcal{M}_3 = 8$ Euros et $r = 10\%$.

Exercice 4

On s'intéresse au comportement d'un client "C" dans le cadre d'un modèle de rétention (en temps discret). Pour tout entier $t \geq 0$ on désigne par p_t la probabilité de "survie" de "C" jusqu'à la fin de la période t et on suppose que $p_0 = 1$ (0 est la période initiale). On note par T "la durée de vie" du client, cette variable aléatoire est en fait égale au nombre de période(s) à la fin desquelles le client est encore en "vie".
1) Montrer que pour tout entier $t \geq 0$, on a $p_t = \mathbb{P}(T \geq t + 1)$.
2) Pour tout entier $t \geq 1$, on appelle probabilité de rétention du client à la période t, la probabilité que le client "survive" jusqu'à la fin de la période t lorsqu'il a "survécu" jusqu'à la fin de la période $t-1$, c'est-à-dire la probabilité conditionnelle $\mathbb{P}(T \geq t + 1/T \geq t)$. On suppose que cette probabilité conditionnelle ne dépend pas de t et on la note par g. On a donc $g = \mathbb{P}(T \geq 2/T \geq 1) = \mathbb{P}(T \geq 3/T \geq 2) = \mathbb{P}(T \geq 4/T \geq 3) = \ldots$.

a) Montrer que $p_1 = g$, $p_2 = g^2$ et plus généralement que $p_t = g^t$ pour tout $\mathbb{N}$.

b) Déterminer la loi de probabilité de T, puis calculer son espérance, sa variance et son écart-type en fonction de g.

Exercice 5

Jean travaille pour un institut de sondage. Dans le cadre d'une enquête, il doit contacter par téléphone des individus pour connaître leur avis sur une émission télévisée. Il dispose de trois listes d'individus, désignées par L_1, L_2 et L_3.

Afin de sélectionner les individus qu'il contactera, Jean décide de procéder de la manière suivante. Pour le premier appel, c'est-à-dire l'appel numéro 0, il prendra au hasard un individu dans la liste L_1. Pour le deuxième appel, c'est-à-dire l'appel numéro 1, Jean choisira d'abord au hasard, de façon équiprobable, l'une des deux listes qui n'auront pas été concernées par l'appel précédent (l'appel numéro 0), puis il prendra au hasard un individu dans la liste retenue. Plus généralement, pour l'appel numéro t (où $t \geq 1$), Jean choisira d'abord au hasard, de façon équiprobable, l'une des deux listes qui n'auront pas été concernées par l'appel précédent (l'appel numéro $t-1$), puis il prendra au hasard un individu dans liste retenue.

Pour tout $t \in \mathbb{N}$, la variable aléatoire X_t désigne la liste qui sera choisie pour l'appel téléphonique numéro t.

1) Sans faire de calculs, expliquer la raison pour laquelle la suite $(X_t)_{t\in\mathbb{N}}$ est une chaîne de Markov homogène.

2) a) Trouver la matrice de transition de $(X_t)_{t\in\mathbb{N}}$ puis tracer son graphe.

b) Déterminer les états qui sont accessibles l'un à partir de l'autre et ceux qui communiquent entre eux. La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ est-elle irréductible ?

3) Montrer que la matrice ligne $\begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix}$ est une probabilité invariante de la chaîne de Markov $(X_t)_{t\in\mathbb{N}}$.

4) Les listes L_1 et L_2 ne comportent aucun individu qui habite à Calais ; en revanche, 20% des individus de la liste L_3 sont des habitants de cette ville. Calculer la probabilité qu'un habitant de Calais soit choisi par Jean pour son deuxième appel téléphonique, c'est-à-dire l'appel numéro 1.

Exercice 6

M. Dupont fait ses courses une fois par semaine dans l'un des trois supermarchés S_1, S_2 et S_3 ; chaque semaine, il choisit aléatoirement le supermarché où il va se rendre. On admet que les choix successifs de M. Dupont au fil des semaines, peuvent être modélisés par une chaîne de Markov homogène $(X_t)_{t\in\mathbb{N}}$ d'espace d'états $\mathcal{E} := \{1, 2, 3\}$ et dont la matrice de transition est :

$$M = \begin{pmatrix} 1/6 & 1/2 & 1/3 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{pmatrix} = 1/6 \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

$X_t = 1$ signifie que la semaine numéro t, M. Dupont fera ses courses dans S_1 ; $X_t = 2$ signifie que la semaine numéro t, M. Dupont fera ses courses dans S_2 ; $X_t = 3$ signifie que la semaine numéro t, M. Dupont fera ses courses dans S_3. La semaine numéro 0 (semaine initiale) est en fait la troisième semaine du mois d'octobre 2021 ; on admet que cette semaine-ci, les 3 choix possibles du supermarché sont équiprobables.

1) a) Déterminer les lois de probabilité de X_0, X_1 et X_2.

b) Quelle est la probabilité que S_1 soit choisi la semaine numéro 1 ?

c) Quelle est la probabilité que S_2 ne soit pas choisi la semaine numéro 2 ?

d) Calculer l'espérance, la variance et l'écart-type de X_1.

2) Tracer le graphe de la chaîne de Markov $(X_t)_{t\in\mathbb{N}}$.

3) Déterminer les états qui sont accessibles l'un à partir de l'autre et ceux qui communiquent entre eux. La chaîne de Markov $(X_t)_{t\in\mathbb{N}}$ est-elle irréductible ?

4) Soit la matrice ligne

$$L = \begin{pmatrix} 11/49 & 21/49 & 17/49 \end{pmatrix} = 1/49 \begin{pmatrix} 11 & 21 & 17 \end{pmatrix}.$$

a) Calculer le produit matriciel LM.
b) Que peut-on alors dire de la matrice L ?
5) Calculer la probabilité de l'événement $A =$"pendant la période des trois semaines ayant pour numéros 0, 1 et 2, M. Dupont fait ses courses uniquement dans le supermarché S_1".

Bibliographie

[1] Berger P.D. and Nasr N.I. (1998) ; Customer lifetime value : Marketing models and applications. *Journal of Interactive Marketing*, 12 (1), 17–30

[2] Blattberg R.C., Gary G. and Jacquelyn S.T. (2001) ; Customer equity. *Boston : Harward Business School Press*

[3] Calciu M. and Salerno F. (2002) ; Customer value modelling : synthesis and extension proposals. *Journal of Targeting, Measurement and Analysis for Marketing*, 11 (2), 124–147

[4] Courtheaux R.J. (1986) ; Database marketing : developing a profitable mailing plan. *Catalog Age, June-July*

[5] Dwyer F.R. (1989) ; Customer lifetime valuation to support marketing decision marketing. *Journal of Direct Marketing*, 9 (1), 79–84

[6] Faure R., Lemaire B. et Picouleau Ch. (7ème édition 2014) ; Précis de recherche opérationnelle méthodes et exercices d'Application. *Dunod*

[7] Jakson B.B. (1985) ; Winning and keeping industrial customer : the dynamics of customer relationships. *Lexington, M.A. : Lexingtion Books*

[8] Kemeny J. G. and Snell J. L. (1976) ; Finite Markov chains. *Springer - Verlag, New York, USA*

[9] Pfeifer P.E. and Carraway R.L. (2000) ; Modeling customer relationships as Markov chains. *Journal of Interactive Marketing*, 14 (2), p. 43–55

[10] Ross Sh.M. (11th edition 2014) ; Introduction to probability models. *Elsevier*

[11] St-Pierre Fortin J. (2018) ; Modélisation par chaînes de Markov de la gestion des relations clients. *HEC Montréal, mémoire présenté en vu de l'obtention du grade de maîtrise ès sciences*