N'aimez-vous pas prévoir le futur parfois?
Chaîne de Markov discrète

1. Introduction

Supposez qu'il y ait un système physique ou mathématique à n états possibles tel qu'à n'importe quel moment, le système soit dans un et seulement un de ses n états. Supposez aussi que durant une période d'observation donnée, disons la $k^{\text{ème}}$ période, la probabilité que le système soit dans un état particulier dépende seulement de son état à la période $k-1$. Un tel système s'appelle une *chaîne de Markov* ou un *processus de Markov*. Dans la suite du texte, les mots soulignés sont définis dans un glossaire en annexe.

Définition 1 : Processus stochastique

On s'intéresse à l'évolution d'un phénomène au cours du temps. Cette évolution est aléatoire, on va la représenter par une famille de variables aléatoires.

Un *processus stochastique* (ou *aléatoire*) est alors : $\{X_t\,;\, t \in T\}$

Où :

- X_t est une variable aléatoire, indexée par le paramètre t, qui est un instant. Les variables aléatoires sont définies sur un espace probabilisé (Ω, E, P). D'une façon générale X_t est l'état du système à l'instant t.
- T est un ensemble d'instants. Il peut prendre deux types de valeurs :
  - fini ou infini mais dénombrable : le *temps* est alors dit *discret*,
  - non dénombrable (T est alors un intervalle de P) : le *temps* est dit *continu*.

X_t prend donc un certain nombre de valeurs au cours du temps. Ainsi $X_t \in S$. L'ensemble S est l'*espace des états*, c'est l'ensemble des états possibles du processus, il peut-être :

- *discret* (c'est-à-dire fini ou infini dénombrable). Le processus est alors appelé une *chaîne*,
- non dénombrable : c'est alors un *processus continu*.

Dans ce dossier, on va travailler avec S fini et T discret.

Définition 2 : Propriété de Markov

Soit une chaîne $\{X_t\,;\, t \geq 0\}$ à temps continu ($T = [0\,;+\infty[$) définie sur l'espace d'états S. On dit qu'elle satisfait à la propriété de Markov quand $\forall t > 0\ \forall h > 0\ \forall n \in \mathrm{N}\ \forall t_1, t_2, \ldots, t_n \in\ ]0\,;t[\ \forall I, s, s_1, \ldots, s_n \in S$, on a :

$$P(X_{t+h} = I \mid X_t = s \text{ et } X_{t_1} = s_1 \text{ et } X_{t_2} = s_2 \text{ et } \cdots \text{ et } X_{t_n} = s_n) = P(X_{t+h} = I \mid X_t = s),$$

où pour deux évènements $A, B \in E$, $P(A \mid B)$ désigne la probabilité conditionnelle de A sachant B. Cela traduit le fait que pour connaître l'évolution future du processus, il suffit de connaître la valeur de son état à l'instant présent. Une telle *chaîne* est dite *de Markov* ou *markovienne*.

Définition 3 : Chaîne de Markov à temps discret

Soit la chaîne $\{X_n\,;\, n \in \mathrm{N}\}$ à temps discret ($T = \mathrm{N}$), $X_n \in S$. Elle est dite *de Markov* quand :

$$\forall n \in \mathrm{N},\ \forall s_0, \ldots, s_n \in S,\ P(X_n = s_n \mid X_0 = s_0 \text{ et } \cdots \text{ et} X_{n-1} = s_{n-1}) = P(X_n = s_n \mid X_{n-1} = s_{n-1})$$

Une chaîne de Markov à temps discret est *homogène dans le temps* si :

$$\forall n > 0,\ \forall k \geq 0,\ \forall s, \sigma \in S,\ P(X_n = s \mid X_{n-1} = \sigma) = P(X_{n+k} = s \mid X_{n+k-1} = \sigma).$$

Définition 4 : Matrice de Transition ou de Passage

Dans la suite on suppose que l'espace des états a r éléments, notés $S = \{1, 2, \dots, r\}$. On s'intéresse aux probabilités $P(X_n = j \mid X_{n-1} = i) = p_{ij}$. On a alors r^2 valeurs, que l'on regroupe dans une matrice :

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

Cette matrice est particulière : toutes les valeurs sont comprises entre 0 et 1 et la somme sur les lignes vaut 1. Cette matrice est aussi appelée **matrice stochastique**.

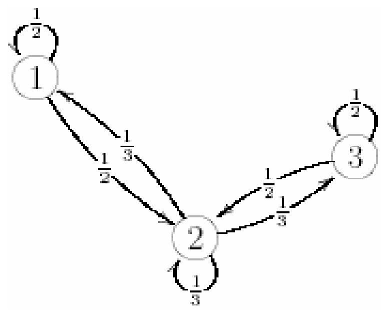
Propriétés 1:

- Toute matrice stochastique $\boldsymbol{P}$ admet le vecteur colonne dont toutes les composantes sont égales à 1 (noté C) comme vecteur propre associé à la valeur propre 1. Réciproquement, si $\forall\ (i, j) \in [1\ ;\ r]^2\ p_{ij} \geq 0$ et $\boldsymbol{P} \times C = C$, alors $\boldsymbol{P}$ est une matrice stochastique.
- Si $\boldsymbol{P}$ est une matrice stochastique, alors pour tout $k \in \mathrm{N}$, $\boldsymbol{P}^k$ est une matrice stochastique.

Exemple :

$$\boldsymbol{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

On valide bien que les sommes des valeurs sur les lignes sont égales à 1. Cette matrice dispose de 3 états. On peut aussi la représenter par un graphe orienté où les sommets sont les états et les arrêtes sont pondérées par les probabilités de passage entre les états :



Le calcul de $\boldsymbol{P}^2$ donne :

$$\boldsymbol{P}^2 = \begin{bmatrix} 5/12 & 5/12 & 1/6 \\ 5/18 & 4/9 & 5/18 \\ 1/6 & 5/12 & 5/12 \end{bmatrix}$$

2. Probabilité de transition en *m* étapes

Pour une **chaîne homogène**, on s'intéresse à la probabilité conditionnelle :

$$P(X_{k+m} = j \mid X_k = i) = P(X_m = j \mid X_0 = i) = p_{ij}^{(m)}.$$

C'est la probabilité de transition de l'état i à l'état j en m étapes. On note $\boldsymbol{P}^{(m)}$ la matrice d'ordre r dont l'élément situé à l'intersection de la ligne i et de la colonne j vaut $p_{ij}^{(m)}$.

Remarque : $\boldsymbol{P}^{(1)} = \boldsymbol{P}$ et $\boldsymbol{P}^{(0)} = \boldsymbol{I}$ (la matrice identité).

Théorème 1 : $\forall m \geq 0$, la probabilité de passage de i à j en m étapes est égale à l'élément (i, j) de la matrice $\boldsymbol{P}^m$.
Preuve. La démonstration se fait par récurrence sur m en appliquant la formule des probabilités totales. ■

Ce résultat est un cas particulier de l'équation de Chapman - Kolmogorov.
Corollaire 1 : Soit $\boldsymbol{P}^{(m)}$ la matrice de transition en m étapes d'une chaîne de Markov, alors $\forall \ell > 0$, $\boldsymbol{P}^{(\ell+m)}$ $\boldsymbol{P}^{(\ell)}\boldsymbol{P}^{(m)}$, ce qui donne :

$$p_{ij}^{(\ell+m)} = \sum_{k=1}^{r} p_{ik}^{(\ell)} p_{kj}^{(m)} .$$

Définition : Période, état apériodique La *période* d d'un état i d'une chaîne de Markov est égale au plus grand diviseur commun de tous les n tels que $p_{ii}^{(n)} > 0$. Si $d > 1$, alors l'état i est *périodique*, sinon il est *apériodique*.

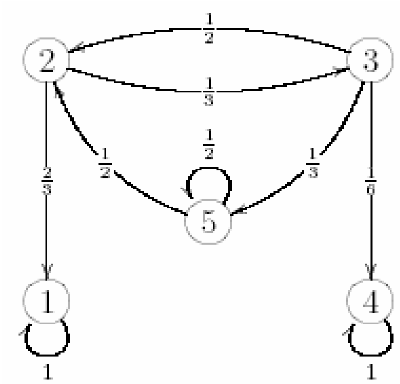
3. Classification

Pour une chaîne de Markov à temps discret homogène, nous dirons que l'état j est *accessible* depuis l'état i quand il existe au moins un chemin de probabilité non nulle menant de i à j. Ce qui s'écrit $\exists n \in \mathrm{N}$, $p_{ij}^{(n)} > 0$. Nous dirons que les états i et j *communiquent* quand l'état j est accessible depuis l'état i et quand l'état i est accessible depuis l'état j. Ce qui s'écrit $\exists n \in \mathrm{N}$, $p_{ij}^{(n)} > 0$ et $\exists m \in \mathrm{N}$, $p_{ji}^{(m)} > 0$. La relation « communiquer » est une relation d'équivalence. On définit alors les classes d'équivalence formées de tous les éléments qui communiquent entre eux, c'est l'opération de *classification*.

Exemple : Soit

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 1/3 & 0 & 0 \\ 0 & 1/2 & 0 & 1/6 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

qui donne le graphe :



La *classification* donne ici trois classes : $C_1 = \{1\}$, $C_2 = \{2, 3, 5\}$ et $C_3 = \{4\}$. A partir de là, on peut écrire le *graphe réduit*, dont les sommets sont les classes et pour lequel un arc, par exemple de C_u vers C_v, signifie qu'il existe $i \in C_u$ et $j \in C_v$ tels que $p_{ij} > 0$.

$$C_1 \longleftarrow C_2 \longrightarrow C_3$$

Définitions 4 : Classes transitoires, persistantes, absorbantes

– Une *classe* est *transitoire* si on peut sortir, sinon elle est *persistante*.
– Une *classe* persistante constituée d'un seul élément est *absorbante*.
– Une *chaîne* est *irréductible* si elle ne comporte qu'une seule classe.
– Une *chaîne* est *absorbante* si toutes ses classes persistantes sont absorbantes.

4. Comportement transitoire

Définition 5 : Distribution des états de la chaîne à l'instant *n*

Soit $p_i^{(n)} = P(X_n = i)$. On regroupe toutes les probabilités dans un vecteur ligne appelé *distribution* après *n* transitions :

$$\pi^{(n)} = (p_1^{(n)}, p_2^{(n)}, \dots, p_r^{(n)}) .$$

Les éléments de ce vecteur sont des probabilités, donc compris entre 0 et 1. La somme vaut 1. On remarque que

$$\pi^{(0)} = \text{la distribution initiale ;}$$
$$\pi^{(1)} = \pi^{(0)} \times \boldsymbol{P} ;$$
$$\pi^{(2)} = \pi^{(1)} \times \boldsymbol{P} = \pi^{(0)} \times \boldsymbol{P}^2 ;\ \pi^{(3)} = \pi^{(2)} \times \boldsymbol{P} = \pi^{(1)} \times \boldsymbol{P}^2 = \pi^{(0)} \times \boldsymbol{P}^3$$

On en déduit donc :

$$\pi^{(n+1)} = \pi^{(0)} \boldsymbol{P}^{n+1} = \pi^{(n)} \boldsymbol{P}$$

La collection des $\pi^{(n)}$ ($n \in$ N) représente le *comportement transitoire*.

5. Comportement asymptotique

On veut connaître la loi de distribution $\pi^{(n)}$ quand $n \to +\infty$. On veut savoir :

– y-a-t'il convergence ?
– dépend-t-elle de la distribution initiale ?
– pour un état persistant, quel va être le temps moyen de séjour dans cet état ?
– pour un état transitoire, quel est le nombre de visites à cet état ?

Définition 6 : Distribution stationnaire

Une *distribution* π est *stationnaire* ou *invariante* si : $\pi = \pi \times \boldsymbol{P}$

Exemple : Pour

$$\boldsymbol{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

la distribution $\pi = (1/2, 1/4, 1/4)$ est invariante.

Remarque : Si la distribution limite existe, elle est invariante.

Nous admettrons les trois résultats suivants.
Propriété 2 : Toute chaîne de Markov finie possède toujours au moins une distribution invariante.

Théorème 2 : Une chaîne de Markov finie possède autant de distributions invariantes linéairement indépendantes que l'ordre de multiplicité de la valeur propre 1 de sa matrice de transition. De plus, cette multiplicité est égale au nombre des classes persistantes de la chaîne.

Propriété 3 : Toute valeur propre λ d'une matrice stochastique vérifie $|\lambda| \leq 1$.

Exemple : Soit la matrice stochastique :

$$\boldsymbol{P} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Son polynôme caractéristique vaut $\Phi(\lambda) = \left(\lambda - \frac{1}{4}\right)(\lambda - 1)^2\left(\lambda + \frac{1}{2}\right)$, on aura donc deux lois stationnaires indépendantes.

On va maintenant chercher les distributions $\pi = (p_1, p_2, p_3, p_4)$ stationnaires : $\pi = \pi \times \boldsymbol{P}$. On obtient un système de 4 équations à 4 inconnues, auquel il faut ajouter la condition $1 = \pi \times C$, c'est-à-dire $p_1 + p_2 + p_3 + p_4 = 1$. On trouve (par exemple) $\pi_1 = (0, 1, 0, 0)$ et $\pi_2 = (0, 0, 2/3, 1/3)$ et tout barycentre $\lambda\pi_1 + (1-\lambda)\pi_2$ (avec $0 \leq \lambda \leq 1$) est une distribution stationnaire.

Théorème 3 : Distribution limite Si $\lim\limits_{n\to+\infty} \pi^{(n)}$ existe et est indépendante de la distribution initiale $\pi^{(0)}$, alors la chaîne possède une seule classe persistante. De plus, $\lim\limits_{n\to+\infty} \pi^{(n)}$ existe et est indépendante de la distribution initiale $\pi^{(0)}$ si et seulement si $\lim\limits_{n\to+\infty} \boldsymbol{P}^n$ existe et est égale à une matrice de transition $\boldsymbol{P}^*$ dont toutes les lignes sont égales entre elles. Une ligne de $\boldsymbol{P}^*$ est alors la distribution limite.

Preuve. Pour montrer la première assertion, supposons qu'il existe au moins deux classes persistantes, notons i et j deux états appartenant à deux classes persistantes distinctes. Si l'on initie le processus avec une distribution dont toutes les probabilités p_k sont nulles sauf pour $k = i$ (resp. $k = j$), alors la distribution limite sera nulle pour tous les états n'appartenant pas à la même classe que i (resp. j). Pour montrer la deuxième assertion, notons π_i la distribution qui est nulle partout sauf en i où elle vaut bien sûr 1 et notons π^* la distribution limite. On a

$$\pi^* = \lim_{n\to\infty} \pi_i \boldsymbol{P}^n = \lim_{n\to\infty} [\boldsymbol{P}^n]_i = [\boldsymbol{P}^*]_i$$

Réciproquement, si $\boldsymbol{P}^*$ existe et si $p_{ij}^* = p_j^*$ pour tout état i, on a $\lim\limits_{n\to\infty} \pi^{(0)}\boldsymbol{P}^n = \pi^{(0)} \lim\limits_{n\to\infty} \boldsymbol{P}^n = \pi^{(0)}\boldsymbol{P}^* = \pi^*$ avec $\pi_j^* = \sum_{i\in S} \pi_i^{(0)} p_j^* = p_j^* \sum_{i\in S} \pi_i^{(0)} = p_j^*$. ■

<u>Théorème 4 : Existence de la distribution limite</u> Soit $\boldsymbol{P}$ la matrice de transition d'une chaîne irréductible et apériodique (c'est-à-dire dont tous les états sont apériodiques). Les propriétés suivantes sont vérifiées :

- pour toute distribution initiale $\pi^{(0)}$: $\lim\limits_{n\to+\infty} \pi^{(n)} = \lim\limits_{n\to+\infty} \pi^{(0)}\boldsymbol{P}^{(n)} = \pi^*$
- π^* est l'unique solution du système $\pi = \pi \times \boldsymbol{P}$ et $1 = \pi \times C$.
- $\forall i \in S$: $\mu_i = \dfrac{1}{\pi_i^*}$ est égal au nombre moyen de transitions entre deux visites successives à l'état *i*.

Cas des chaînes réductibles

On suppose que l'on a plusieurs classes. On a :

- Des classes persistantes (ou absorbantes) ;
- Des classes transitoires.

<u>Théorème 5 :</u> Pour tout état initial *i*, la probabilité d'être dans un état persistant à l'étape *n* tend vers 1 quand *n* tend vers l'infini. De la même façon, la probabilité d'être dans un état transitoire tend vers 0.

6. Forme canonique de la matrice de transition *P*

On donne une forme canonique à $\boldsymbol{P}$ en renumérotant les états : on place les états des classes persistantes en premier, puis les états transitoires.

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_2 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{P}_k & \boldsymbol{0} \\ \boldsymbol{R}_1 & \boldsymbol{R}_2 & \cdots & \boldsymbol{R}_k & \boldsymbol{Q} \end{pmatrix}$$

Chaque $\boldsymbol{P}_i$ est une chaîne de Markov irréductible sur la classe persistante C_i.

Si la chaîne de Markov est absorbante, la forme canonique de $\boldsymbol{P}$ sera :

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{Q} \end{pmatrix}. \tag{1}$$

Si on calcule $\boldsymbol{P}^2$ ou plus généralement $\boldsymbol{P}^n$ on obtient :

$$\boldsymbol{P}^2 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{R}+\boldsymbol{QR} & \boldsymbol{Q}^2 \end{pmatrix} \text{ et } \boldsymbol{P}^n = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \sum\limits_{i=0}^{n-1} \boldsymbol{Q}^i\boldsymbol{R} & \boldsymbol{Q}^n \end{pmatrix}, \text{ donc } \lim_{n\to+\infty} \boldsymbol{P}^n = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ (\boldsymbol{I}-\boldsymbol{Q})^{-1}\boldsymbol{R} & \boldsymbol{0} \end{pmatrix}.$$

On note $\boldsymbol{N} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}$, la matrice *fondamentale* de la chaîne de Markov. Le résultat que l'on vient de voir va nous permettre de calculer π^*.

<u>Théorème 6 :</u> Soit une chaîne de Markov débutant de l'état transitoire *i*. Le nombre moyen de périodes passées dans un état transitoire *j* est égal à l'élément n_{ij} de la matrice $\boldsymbol{N}$. De plus, partant d'un état transitoire *i*, le nombre moyen de transitions avant d'atteindre un état absorbant est égal à la somme des éléments de la $i^{\text{ème}}$ ligne de $\boldsymbol{N}$.

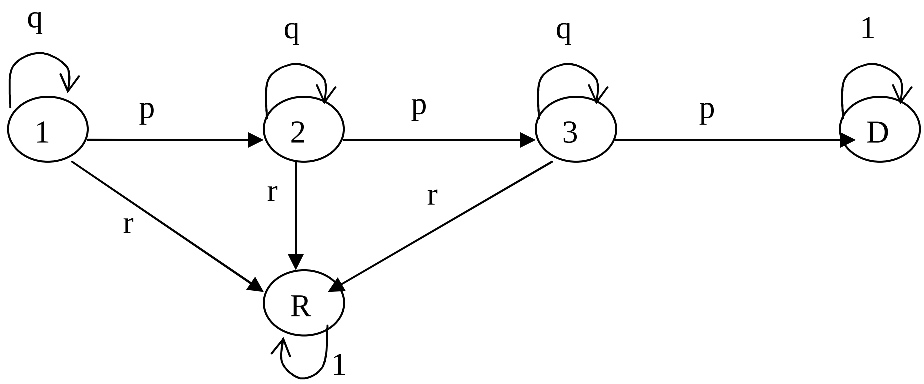
Preuve. Pour $n > 0$, considérons le nombre moyen de visites à l'état j pendant n périodes, partant de l'état i, où i et j sont transitoires. Remarquons que i et j étant tous les deux transitoires, l'élément p_{ij} de la matrice $\boldsymbol{P}$ est localisé dans le bloc $\boldsymbol{Q}$ de la relation (1) et plus généralement, l'élément $p_{ij}^{(k)}$ de la matrice $\boldsymbol{P}^k$ est localisé dans le bloc $\boldsymbol{Q}^k$. Pour l'état initial i, considérons les n variables aléatoires de Bernoulli $B_k = 1$ si $X_k = j$, sinon $B_k = 0$ pour $1 \leq k \leq n$. Le nombre de visites à l'état j pendant n périodes vaut $B_1 + \cdots + B_n$ et sa valeur moyenne vaut

$$\mathrm{E}\left[\sum_{k=1}^{n} B_k \,\middle|\, X_0 = i\right] = \sum_{k=1}^{n} \mathrm{E}\left[B_k \mid X_0 = i\right] = \sum_{k=1}^{n} p_{ij}^{(k)} = \sum_{k=1}^{n} q_{ij}^{(k)} \xrightarrow[n\to\infty]{} n_{ij} .$$

La deuxième assertion s'en déduit. ■

7. Etudes dans une Grande Ecole :

Les études dans une Grande Ecole durent trois ans, à l'issue de chaque année, chaque élève a une probabilité p de passer dans l'année supérieure (ou d'obtenir son diplôme s'il est en troisième année) ; une probabilité q de redoubler, r d'être renvoyé. Le cursus d'un élève peut être modélisé par une chaîne de Markov à cinq états dont deux sont absorbants. Le graphe associé à la chaîne de Markov est le suivant :



avec R et D correspondant respectivement au Renvoi de l'école et à l'obtention du Diplôme.

A. Probabilité d'obtention du diplôme

L'évaluation de la probabilité a_{iD} qu'un élève obtienne son diplôme étant dans l'état i (i = 1, 2, ou 3) peut être obtenue en décomposant l'événement {partant dans l'état i, le système va être absorbé par D} selon l'issue de la première transition et en utilisant le caractère sans mémoire de l'évolution de système. Il vient :

$$a_{1D} = q \,.\, a_{1D} + p \,.\, a_{2D} + r \,.\, a_{RD}$$
$$a_{2D} = q \,.\, a_{2D} + p \,.\, a_{3D} + r \,.\, a_{RD}$$
$$a_{3D} = q \,.\, a_{3D} + p \,.\, a_{DD} + r \,.\, a_{RD}$$

Mais nous avons $a_{RD} = 0$, $a_{DD} = 1$ et $p + q + r = 1$, d'où le nouveau système linéaire à résoudre :

$$(p + r) \,.\, a_{1D} - p \,.\, a_{2D} = 0$$
$$(p + r) \,.\, a_{2D} - p \,.\, a_{3D} = 0$$
$$(p + r) \,.\, a_{3D} = p$$

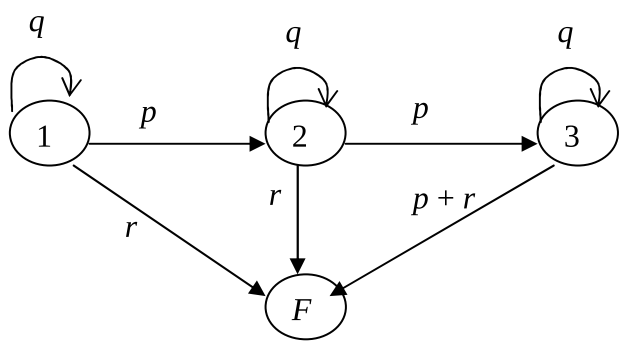
dont la solution est :

$$a_{3D} = \frac{p}{p+r} \qquad a_{2D} = \left(\frac{p}{p+r}\right)^2 \qquad a_{1D} = \left(\frac{p}{p+r}\right)^3$$

Pour tout état i, nous avons $a_{iD} + a_{iR} = 1$ (tout élève finira soit par être renvoyé, soit par obtenir son diplôme) ; on en déduit les valeurs de a_{1R}, a_{2R} et a_{3R}.

B. Le calcul de la durée moyenne des études

La fin des études est associée aux états R ou D. Ces deux états peuvent être regroupés en un état F et nous considérons alors la chaîne ci-dessous à quatre états, dont seul l'état F est absorbant avec $p_{3F} = p + r$.



Appelons $\overline{t_i}$, pour i = 1, 2, 3, le temps moyen de séjour dans les états transitoires (1, 2, 3) avant absorption par F (la durée moyenne des études est alors $\overline{t_1}$) ; en considérant le résultat de la première transition et en utilisant le caractère sans mémoire de la chaîne (le temps moyen d'absorption à partir de l'état i ne dépend que de i), nous pouvons écrire :
$\overline{t_1} = r.1 + q.(1+\overline{t_1}) + p.(1+\overline{t_2})$, c'est-à-dire :

$$\overline{t_1} = 1 + q.\overline{t_1} + p.\overline{t_2}$$

De même :

$$\overline{t_2} = 1 + q.\overline{t_2} + p.\overline{t_3} \qquad \text{et} \qquad \overline{t_3} = 1 + q.\overline{t_3}$$

La solution de ce système est :

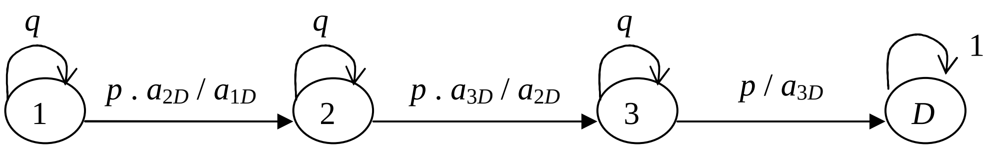
$$\overline{t_3} = \frac{1}{1-q} \qquad \overline{t_2} = \frac{2.p+r}{(1-q)^2} = \frac{2.p+r}{(p+r)^2} \qquad \overline{t_1} = \frac{3.p^2 + 3.r.p + r^2}{(p+r)^3}$$

C. Calcul de la durée moyenne d'obtention d'un diplôme

Il faudra s'intéresser aux seules trajectoires qui se terminent par D, c'est-à-dire que l'on s'intéresse au cursus d'un élève dont on sait qu'il obtiendra son diplôme. Les probabilités de transition $p_{i,j}$ doivent alors être remplacées par les probabilités de passage de i à j conditionnées par une absorption par l'état D. En appliquant la définition des probabilités conditionnelles, nous avons :

P[passage de i à j | absorption par D] = P[passage de i à j et absorption par D] / P[absorption par D] = $p_{ij} \,.\, a_{jD} / a_{iD}$.

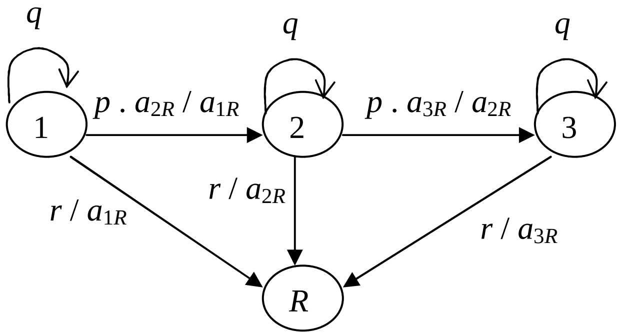
Pour les "heureux" élèves, la chaîne est alors la suivante :



Il est alors simple de calculer par la même méthode qu'en 7.B. le temps moyen d'absorption par D, on trouve :

$$\overline{t_1^D} = \frac{3}{p+r} \qquad \overline{t_2^D} = \frac{2}{p+r} \qquad \overline{t_3^D} = \frac{1}{p+r}$$

Considérons désormais les cursus des élèves « renvoyés». Pour eux les probabilités de passage doivent être conditionnées par l'absorption par l'état R, p_{ij} devient $p_{ij}\ .\ a_{jR}\ /\ a_{iR}$. La chaîne associée est alors la suivante :



Le calcul des temps d'absorption par l'état R donne, en employant la même méthode qu'en 7.B :

$$\overline{t_3^R} = \frac{1}{p+r} \qquad \overline{t_2^R} = \frac{1}{p+r} \times \frac{r+3.p}{r+2.p} \qquad \overline{t_1^R} = \frac{1}{p+r} \times \frac{6.p^2 + r^2 + 4.p..r}{r^2 + 3.p^2 + 3.r.p}$$

Il est alors facile de vérifier que le temps moyen de fin d'études $\overline{t_1}$ doit satisfaire la relation suivante :

$$\overline{t_1} = \overline{t_1^D}.\ a_{1D} + \ \overline{t_1^R}.\ a_{1R}$$

Annexe

Glossaire :

<u>Espace probabilisé</u> : un espace probabilisé modélise une <u>expérience aléatoire</u> ; c'est la donnée d'un triplet (Ω, E, P) où Ω est un ensemble dont les éléments sont les résultats possibles de l'expérience aléatoire, E est un sous-ensemble de $\Pi(\Omega)$ --- ensemble des parties de Ω----, dont les éléments sont appelés évènements et qui satisfait aux trois axiomes :

- $\phi \in$ E ;
- Si $A \in$ E, alors le complémentaire de A dans Ω, noté A^c, est aussi dans E ;
- Poute famille finie ou infinie dénombrable d'éléments de E, leur réunion est encore dans E.

Autrement dit, E est stable par passage au complémentaire et par réunion finie ou infinie dénombrable. Le troisième élément du triplet est une application, appelée probabilité, $P : A \in$ E # $P(A) \in [0 ; 1]$ qui vérifie

- $P\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} P(A_i)$ pour toute famille finie ou infinie dénombrable $(A_i)_{i\in I}$ d'évènements deux à deux disjoints.

$P(A)$ est la probabilité d'occurrence de l'évènement A.

<u>Evénement :</u> voir espace probabilisé.

<u>Expérience aléatoire :</u> une expérience, c'est-à-dire l'observation d'un phénomène, reproductible dans le temps et dans l'espace est dite aléatoire quand, une fois fixées les conditions expérimentales, l'expérience ne donne pas toujours le même résultat, mais un résultat parmi un ensemble de résultats possibles. Par exemple, le lancer d'un dé.

<u>Formule des probabilités totales</u> : soient $A_1, \ldots, A_n \in$ E, n événements formant une partition de Ω, $\forall B \in$ E, on a

$$P(B) = \sum_{i=1}^{n} P(B|\, A_i) P(A_i).$$

<u>Probabilité conditionnelle :</u> Pour deux évènements A, $B \in$ E, tels que $P(B) > 0$, la probabilité conditionnelle de A sachant B, notée $P(A \mid B)$ est définie par

$$P(A|\, B) = \begin{cases} P(A \cap B)/P(B) & \text{si } P(B) > 0 \\ 0 & \text{sinon} \end{cases}$$

On montre que l'application $P(. \mid B)$: $A \in$ E # $P(A \mid B)$ est une probabilité.

<u>Variable aléatoire</u> : une application X : $\omega \in \Omega$ # $X(\omega) \in$ P est un variable aléatoire quand pour tout réel $a \in$ P, $X^{-1}(]-\infty ; a]) = \{\omega \in \Omega |\, X(\omega) \le a\}$ est un évènement (c'est-à-dire un élément de E).