Jacinthe Rose[1] est inquiète : son fils a commis un délit et elle craint qu'il ne fasse une carrière de criminel. Elle consulte madame Irma, la voyante, pour connaître l'avenir de son fils.

Irma
Non, les feuilles de thé c'est complétement dépassé. On utilise maintenant les probabilités et les chaînes de Markov.

Jacinthe
C'est quand même très compliqué une carrière criminelle!

Récidivisme

Marc Laforest
École Polytechnique
de Montréal
André Ross
Cégep de Lévis-Lauzon

Jacinthe Rose
Madame Irma, mon fils a commis un délit. J'ai peur qu'il ne devienne un criminel. Pouvez-vous me rassurer ?

Irma
Je peux établir la longueur moyenne $\overline{L}$ d'une carrière criminelle en fonction de la probabilité r qu'un criminel récidive, mais il faut connaître la probabilité que votre fils récidive pour pouvoir connaître son avenir.

Jacinthe
Probabilité? Vous ne lisez plus dans les feuilles de thé?

Irma
On peut simplifier l'analyse en considérant quatre états mesurables dans une carrière criminelle :

1. est un citoyen honnête,
2. commet un crime,
3. est en état d'arrestation,
4. est en prison.

Votre fils a commis un délit mais si, dans toutes ses actions futures, il se comporte en citoyen honnête, sa carrière criminelle peut se décrire par le graphique suivant :

$$\begin{array}{ccccccccccccc} E_1 & & E_2 & & E_3 & & E_4 & & E_5 & & E_6 & & \\ 2 & \rightarrow & 1 & \rightarrow & 1 & \rightarrow & 1 & \rightarrow & 1 & \rightarrow & 1 & \rightarrow & \cdots \end{array}$$

Jacinthe
Je souhaite tellement que ce soit le cas.

Irma
Je vous le souhaite également. Supposons maintenant qu'après ce premier délit, il est arrêté, puis incarcéré.

Jacinthe
Vous voulez me faire mourir!

Irma
Supposons également qu'à sa libération il commet un autre délit puis devient citoyen honnête. On pourrait alors illustrer sa carrière de criminel par le diagramme suivant :

$$\begin{array}{ccccccccccccc} E_1 & & E_2 & & E_3 & & E_4 & & E_5 & & E_6 & & \\ 2 & \rightarrow & 3 & \rightarrow & 4 & \rightarrow & 2 & \rightarrow & 1 & \rightarrow & 1 & \rightarrow & \cdots \end{array}$$

1. *Référence à Mr. Pink du film "Reservoir Dogs".*

Reconnaissez-vous que l'on peut décrire ainsi toutes les carrières criminelles possibles?

Jacinthe
Oui, mais je ne vois pas comment vous allez faire votre prédiction.

Irma
En fait, on peut alléger la représentation en disant qu'une carrière criminelle est une suite d'états

$$\mathbf{2} \to E_2 \to E_3 \to E_4 \to \cdots$$

où les E_i représentent les états en cours de carrière, les indices indiquant la succession des états, et :

$$E_i \in \{\mathbf{1, 2, 3, 4}\}$$

L'ensemble de tous les choix de carrières criminelles est donc équivalent à l'ensemble de tous les parcours qui commencent à l'état **2** et qui sont formés des transitions admissibles entre les états. L'ensemble des carrières criminelles s'écrit donc :

$$\Omega = \{ E_1 \to E_2 \to E_3 \to E_4 \to \cdots \}$$
$$\text{où } E_1 = 2 \text{ et } E_i \in \{\mathbf{1, 2, 3, 4}\}.$$

Jacinthe
Cela donne beaucoup de suites.

Irma
En effet. Pour aller plus loin dans ma prévision, j'utilise l'hypothèse de Markov selon laquelle les états passés n'ont pas d'influence sur les états futurs.

Jacinthe
Vous voulez dire que les criminels ne se souviennent pas d'avoir été emprisonnés et que la probabilité de commettre un délit une deuxième fois est la même que la première fois?

Irma
Pas pour chacun pris individuellement, mais pour prévoir le comportement général des criminels, c'est une hypothèse raisonnable qui nous permet de considérer que les événements sont indépendants.

Jacinthe
Quel est l'intérêt de considérer l'indépendance des événements?

Irma
Lorsque deux événements sont indépendants, alors la probabilité qu'ils se produisent tous les deux est le produit des probabilités de chacun.

Jacinthe
Je ne comprends pas.

Irma
Supposons que l'on note :

- r, la probabilité qu'une personne récidive,
- p, la probabilité qu'une personne ne soit pas arrêtée si elle a commis un délit,
- q, la probabilité qu'une personne arrêtée pour un délit ne soit pas condamnée à la prison.

L'indépendance des événements fait que la probabilité de la séquence

$$\mathbf{2} \to \mathbf{3} \to \mathbf{4}$$

est :

$$P(\mathbf{2} \to \mathbf{3} \to \mathbf{4}) = P(\mathbf{2} \to \mathbf{3}) \cdot P(\mathbf{3} \to \mathbf{4})$$
$$= (1-p)\cdot(1-q)$$

Jacinthe
Que signifient les facteurs $(1 - p)$ et $(1 - q)$?

Irma
Si la probabilité de ne pas être arrêté est p, la probabilité d'être arrêté est $(1 - p)$.

Jacinthe
Je vois, mais une carrière criminelle comporte plus d'états que cette simple séquence.

Irma
C'est vrai, mais l'hypothèse de Markov permet de considérer que la probabilité d'une carrière est donnée par le produit des probabilités des changements d'état qui constituent cette carrière.

Je vous donne un autre exemple. Si on considère la carrière criminelle

$$\begin{array}{ccccccccccccc} E_1 & & E_2 & & E_3 & & E_4 & & E_5 & & E_6 & & \\ \mathbf{2} & \to & \mathbf{3} & \to & \mathbf{4} & \to & \mathbf{2} & \to & \mathbf{1} & \to & \mathbf{1} & \to & \cdots \end{array}$$

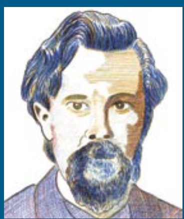
La probabilité de cette carrière est :

$$P(\mathbf{2} \to \mathbf{3} \to \mathbf{4} \to \mathbf{2} \to \mathbf{1} \ldots)$$
$$= P(\mathbf{2} \to \mathbf{3}) \cdot P(\mathbf{3} \to \mathbf{4}) \cdot P(\mathbf{4} \to \mathbf{2}) \cdot P(\mathbf{2} \to \mathbf{1}) \ldots$$
$$= (1-p)\cdot(1-q)\cdot r \cdot [p \cdot (1-r)] \cdot 1 \cdot 1 \cdot 1 \ldots$$

Jacinthe
Que signifie la suite de 1 à la fin?

Irma
La suite de 1 tient compte d'une hypothèse additionnelle : lorsqu'un criminel est devenu honnête, il le reste jusqu'à la fin de ses jours. Dans une chaîne de Markov, on dit que **1** est un état absorbant et toute chaîne de Markov ayant un état absorbant se stabilise dans cet état.

Andreï Andreïevich Markov (1856-1922)



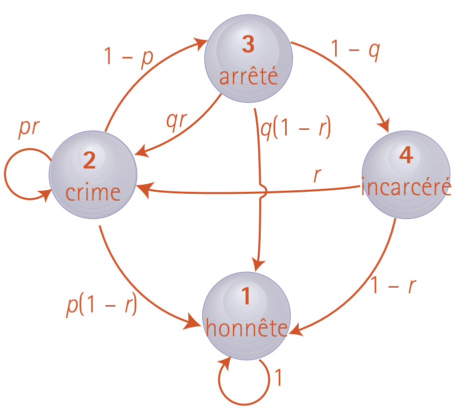
Mathématicien russe, il étudia les mathématiques à l'Université de Saint-Pétersbourg où il fut l'élève de Pafnuty Tchebychev. Il termina ses études en 1878 et enseigna lui-même à Saint-Pétersbourg à partir de 1886. Malgré ses nombreux travaux sur la théorie des nombres, les fractions continues et les équations différentielles, sa renommée est attribuable surtout à ses travaux en probabilité et, en particulier, aux chaînes qui portent son nom. Celles-ci sont appliquées en finance, en sciences sociales, en mécanique quantique, en génétique et en physique atomique.

Ces chaînes permettent d'étudier des situations mettant en cause des variables aléatoires dont l'état futur dépend de l'état présent, mais est indépendant de la façon dont l'état présent est intervenu à partir des états précédents. Ces travaux ont marqué le début de l'étude des *processus stochastiques*. Les travaux de Markov ont été poursuivis par Andreï Kolmogorov, Norbert Wiener et Kiyoshi Itō entre autres.

Jacinthe
Cela veut dire qu'à long terme tout criminel devient honnête. Vous trouvez cela réaliste?

Irma
À condition de ne pas mourir avant, bien sûr. C'est parfois mieux d'avoir un modèle simple avec des hypothèses simples qu'un modèle complexe avec beaucoup de petites hypothèses réalistes.

Mais je continue. En tenant compte de l'hypothèse de Markov et du fait que devenu honnête un criminel le demeure, on peut calculer les probabilités de transitions. Ainsi, votre fils a commis un crime; il est dans l'état 2, alors :



- il peut être arrêté avec probabilité $1 - p$,
- il peut récidiver avec une probabilité $p \cdot r$,
- il peut abandonner sa vie criminelle avec probabilité $p \cdot (1 - r)$.

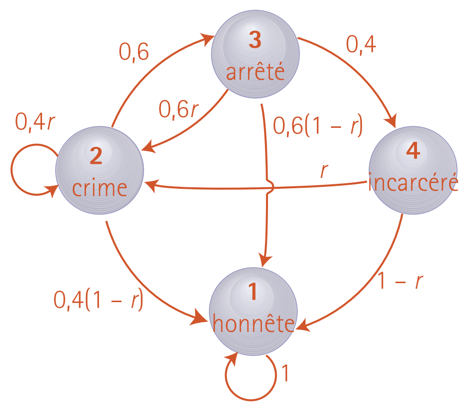
On peut, de la même façon, calculer les probabilités de transition partant de l'état **3** ou de l'état **4**. On peut représenter cela par un *diagramme de transition*.

Jacinthe
Comment faites-vous pour connaître les probabilités p, q et r?

Irma
Pour les probabilités p et q, j'utilise les dernières données statistiques disponibles. Ainsi, en 1998, au Canada 60 % des crimes ont été résolus. La probabilité de ne pas être arrêté après avoir commis un délit est donc de 40 % ($1 - p = 0{,}6$ et $p = 0{,}4$).

Selon le même rapport statistique, 40 % des criminels arrêtés ont été condamnés à la prison. La probabilité d'aller en prison après avoir été arrêté est donc de 40 % ($1 - q = 0{,}4$ et $q = 0{,}6$) .

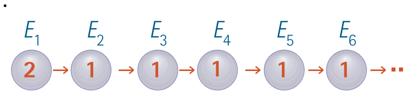
En tenant compte de ces probabilités et de l'indépendance des événements, je peux alors préciser le diagramme de transition.



Jacinthe
Je comprends pourquoi il faut connaître la probabilité que mon fils récidive pour prévoir sa carrière. C'est la seule inconnue du diagramme.

Irma
Exactement. On peut maintenant aborder la notion de *longueur d'une carrière criminelle*. Puisque notre chaîne de Markov a un état absorbant et qu'un jour ou l'autre tout criminel finit par devenir honnête, on peut définir la longueur d'une carrière criminelle. C'est le nombre de transitions avant de devenir honnête.

Par exemple, votre fils a commis un délit. Si, après ce délit, il devient honnête, sa carrière est :



La longueur de sa carrière C est :

$$L(C) = 1$$

puisqu'il y a une seule transition avant qu'il ne devienne honnête.

Par ailleurs si sa carrière est plutôt la suivante :

$$\overset{E_1}{2} \to \overset{E_2}{3} \to \overset{E_3}{4} \to \overset{E_4}{2} \to \overset{E_5}{1} \to \overset{E_6}{1} \to \cdots$$

Dans ce cas, la longueur de sa carrière est 4.

Jacinthe
Vous avez dit pouvoir établir la longueur moyenne $\bar{L}$ d'une carrière criminelle en fonction de la probabilité r. Comment faites-vous?

Irma
Le principe est simple, il faut utiliser un générateur de carrières qui détermine un grand nombre de carrières en tenant compte de la probabilité de récidive. On peut alors déterminer la longueur de chacune de ces carrières.

Jacinthe
Je ne saisis pas.

Irma
Par exemple, si, pour une valeur de r donnée, on génère 10 000 carrières et que parmi celles-ci, la carrière C_i se répète 600 fois, on peut alors estimer que, pour cette valeur de r, la probabilité qu'un criminel ait cette carrière est :

$$P(C_i) = \frac{600}{10\,000} = 6\,\%.$$

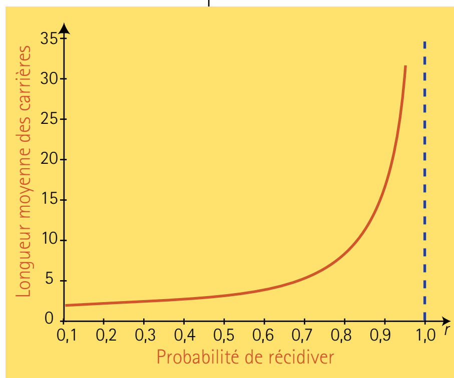
Pour calculer la longueur moyenne des carrières criminelles, on fait la somme suivante :

$$P(C_1) \cdot L(C_1) + P(C_2) \cdot L(C_2) + \cdots$$

Chaque longueur de carrière est multipliée par sa probabilité et cette somme est la longueur moyenne cherchée. On peut écrire cela de façon condensée, soit :

$$\bar{L} = \sum_{C \in \Omega} P(C) \cdot L(C).$$

Pour chaque valeur de r entre 0,1 et 0,95, j'ai fait produire 10 000 carrières par mon ordinateur et cela m'a donné une courbe surprenante.



Jacinthe
On voit que la pente de la courbe augmente très rapidement quand la probabilité de récidive s'approche de 1.

Irma
En effet, une diminution de $r = 0{,}8$ à $r = 0{,}6$, soit une diminution de 20 %, entraîne une baisse de la longueur moyenne des carrières d'environ 8,3 à 4,1, soit une diminution de plus de 50 %.

Jacinthe
Vos explications suggèrent qu'il est très profitable, tant du point de vue social qu'économique, d'encourager les criminels à se réformer.

Irma
C'est en effet la conclusion que l'on peut tirer socialement. Et pour votre garçon : mes feuilles de thé et les prévisions du modèle ne changeront rien!

Jacinthe
Tout n'est donc pas perdu pour mon fils! Je peux et je dois l'aider!

Irma
À lui de travailler à mettre toutes les chances de son côté.

Générateur aléatoire de carrières

On utilise un générateur de nombres aléatoires uniformément distribués dans l'intervalle [0, 1]. L'état initial est $E_1 = \mathbf{2}$. À la première étape, on fait tirer au hasard un nombre θ_1 dans l'intervalle [0, 1] et on choisit l'état E_2 de la façon suivante :

$$E_2 = \begin{cases} \mathbf{1} & \text{si } \theta_1 \in [0,\, p(1-r)[\\ \mathbf{2} & \text{si } \theta_1 \in [p(1-r),\, p[\\ \mathbf{3} & \text{si } \theta_1 \in [p,\, 1] \end{cases}$$

En appliquant cet algorithme pour obtenir un très grand nombre de paires $\mathbf{2} \to E_2$, alors celles-ci seront représentatives des probabilités de transitions a_{21}, a_{22} et a_{23}, où a_{ji} est la probabilité de passer de l'état $\mathbf{j}$ à l'état $\mathbf{i}$.

Supposons que l'on a déjà obtenu les n premiers états de la carrière de M. Rose,

$$C = 2 \to E_2 \to E_3 \to \ldots \to E_n$$

de manière à satisfaire, au moins pour ces états :

$$P(C) = \frac{\text{nombre de fois que } C \text{ est généré}}{\text{nombre de carrières générées}}.$$

Pour déterminer E_{n+1} quand $E_n = \mathbf{j}$, nous tirerons au hasard un nombre $\theta_n \in [0, 1]$ et nous poserons :

$$E_{n+1} = \begin{cases} \mathbf{1} & \text{si } \theta_n \in [0,\, a_{j1}[\\ \mathbf{2} & \text{si } \theta_n \in [a_{j1},\, a_{j1} + a_{j2}[\\ \mathbf{3} & \text{si } \theta_n \in [a_{j1} + a_{j2},\, a_{j1} + a_{j2} + a_{j3}[\\ \mathbf{4} & \text{si } \theta_n \in [a_{j1} + a_{j2} + a_{j3},\, 1] \end{cases}$$