

Chapitre 10 : Probabilités

ECE3 Lycée Carnot

15 décembre 2010

Introduction via quelques exemples

Le concept de probabilité est a priori relativement intuitif : rien de suprenant à ce qu'un dé à six faces normalement constitué tombe en moyenne une fois sur six sur chacune de ses faces (il s'agit toutefois d'un résultat statistique, qui ne garantit par exemple en aucun cas qu'au bout de six lancers on aura obtenu chacun des six résultats possibles). Les probabilités étudiées au lycée restent la plupart du temps dans ce cadre : nombre fini de possibilités, probabilité égale pour chacun des cas possibles, mais en fait, l'étude des probabilités en mathématiques peut se faire dans un cadre beaucoup plus large.

Premier exemple

Reprenons un exemple qui s'inscrit bien dans le cadre vu au lycée : on lance simultanément deux dés à six faces et on note la somme des deux résultats obtenus. Il est assez facile de se convaincre que tous les résultats possibles (en l'occurrence les entiers compris entre 2 et 12) n'apparaîtront pas avec la même fréquence, car il existe par exemple 4 façons d'obtenir une somme égale à 5 ($1 + 4$; $2 + 3$; $3 + 2$ et $4 + 1$), mais une seule d'obtenir 2 (les deux dés doivent tomber sur 1). Pour préciser cela, on peut considérer les choses de la façon suivante : il y a 6 résultats possibles pour le premier dé, autant pour le second, soit un total de 36 possibilités. On obtiendra donc une somme égale à 2 en moyenne une fois sur 36, mais une somme égale à 5 quatre fois sur 36, soit une fois sur 9. Cet exemple illustre bien la nécessité de bien définir quel est l'ensemble de résultats sur lequel on veut travailler, et surtout de vérifier si ces résultats sont équiprobables ou non.

Deuxième exemple

On lance cette fois-ci un seul dé plusieurs fois de suite, jusqu'à ce que le dé tombe sur la face numéro 6. La situation est beaucoup plus compliquée puisqu'il y a ici une infinité de résultats possibles : un résultat où on ne tire qu'un fois le dé et on tombe sur 6, cinq résultats à deux lancers (1 puis 6, 2 puis 6 etc), 36 résultats à trois lancers etc... Déterminer la probabilité d'avoir besoin d'attendre exactement 10 lancers pour obtenir notre premier 6 reste toutefois relativement élémentaire : il faut que chacun des neuf premiers lancers donne un autre résultat que 6, et que le dixième tombe sur 6, soit une probabilité de $\left(\frac{5}{6}\right)^9 \times \frac{1}{6}$. Mais on peut se poser des questions plus complexes à propos de cette expérience : est-il possible de ne jamais obtenir de 6, même après une infinité de lancers? La réponse mathématique est un peu surprenante : oui, c'est possible, mais la probabilité que ça arrive est nulle! Autre question intéressante : combien de lancers faudra-t-il en moyenne pour obtenir notre premier 6? Dans le cas d'un nombre fini de résultats possibles, une moyenne se calcule en faisant la moyenne des résultats possibles pondérés par leurs probabilités respectives. Ici, bien que l'ensemble des résultats soit infini, le même calcul reste possible, il va simplement s'agir désormais d'un calcul de somme de série. Pour les curieux, on constate que la probabilité d'obtenir notre premier 6 au tirage numéro k vaut $\left(\frac{5}{6}\right)^{k-1} \times \frac{1}{6}$, et la moyenne se calcule

via $\sum_{k=1}^{+\infty} k \times \left(\frac{5}{6}\right)^{k-1} \times \frac{1}{6} = \frac{1}{6} \times \frac{1}{\left(1 - \frac{5}{6}\right)^2} = 6$ (simple calcul de somme de série géométrique dérivée).

Il faut donc en moyenne six lancers avant d'obtenir un 6.

Troisième exemple

Une cible de jeu de fléchettes est constituée de trois disques concentriques de rayons respectifs 1, 2 et 3. Un joueur lance une fléchette dans la cible. On suppose (ce n'est pas très réaliste) que la fléchette atteint toujours la cible et que le point atteint dans la cible est aléatoire (avec une probabilité uniforme). On sort manifestement du cadre habituel : le nombre de résultats possibles n'est pas fini, loin s'en faut. On peut tout de même attribuer de manière assez intuitive des probabilités à certains événements : par exemple, il paraît naturel de dire que la probabilité de tomber dans le disque central vaut un neuvième (rapport entre l'aire du disque central et celle de la cible). Mais que dire de l'événement « La fléchette tombe sur un point qui est à une distance rationnelle du centre » (oui, certes, personne ne se pose ce genre de question) ? Pas moyen de calculer facilement l'aire d'une telle chose. On admettra en fait que, dans ce genre de cas, on ne peut calculer les probabilités de tout et n'importe quoi, et qu'il faudra donc choisir quels sont les événements autorisés (cela nous mènera au concept de tribu).

Quatrième exemple

On cherche à étudier une file d'attente (à la Poste, par exemple). À tout moment, il existe une certaine probabilité qu'une personne vienne s'ajouter à la file existante, et chaque personne passe au guichet un temps aléatoire. Ce temps est en pratique borné, mais peut prendre à peu près n'importe quelle valeur positive dans les limites du raisonnable. On pourrait naturellement décider de découper le temps en une multitude de petits intervalles (d'une seconde chacun, par exemple) et se ramener à des probabilités sur un ensemble fini, mais les calculs seraient affreusement lourds. Il est en fait plus logique d'accepter de faire des probabilités sur l'ensemble \mathbb{R}_+ (même si, comme dans le cas précédent, on ne pourra calculer la probabilité de n'importe quoi), et de développer une théorie qui englobera également le cas des probabilités finies. Nous allons nous y atteler de ce pas. Dans ce dernier cas, on rentre dans un domaine des probabilités (les probabilités continues), que vous étudierez plus intensivement l'an prochain, et qui fait intervenir beaucoup de calcul intégral (eh oui...).

1 Vocabulaire

1.1 Expérience aléatoire

Définition 1. Une **expérience aléatoire** est un phénomène ayant des résultats numériques dépendant du hasard.

Exemple : Nous reprendrons tout au long de ce chapitre un exemple particulier pour illustrer les diverses définitions, celui consistant à tirer simultanément cinq cartes au hasard dans un jeu de 32 cartes.

Remarque 1. Insistons une fois de plus sur le fait qu'une expérience aléatoire est par définition non déterministe. L'étude des probabilités permet de faire des prévisions statistiques, mais en aucun cas de prévoir le résultat d'une expérience précise. Autrement dit, vous aurez beau être très fort en probas, ça ne vous aidera pas à décrocher la cagnotte au Loto.

Définition 2. On appelle **univers**, et on note Ω , l'ensemble des résultats possibles d'une expérience aléatoire.

Exemple : Dans notre exemple, Ω est beaucoup trop gros pour qu'on puisse faire la liste de ses éléments, mais on sait par contre que $|\Omega| = \binom{32}{5}$. Attention toutefois à ne pas confondre Ω , qui est un ensemble, et son cardinal, qui est un nombre.

1.2 Événements

Définition 3. Un **événement** (souvent noté A, B, \dots) est un sous-ensemble de l'univers Ω . On dit qu'un événement A est réalisé si le résultat de l'expérience appartient à ce sous-ensemble.

Exemple : En pratique, un événement est la plupart du temps décrit par une propriété plutôt que comme un sous-ensemble. Par abus de langage on dira ainsi qu'on considère l'événement A : « on tire les quatre As et le Roi de pique ». Ce qui nous intéressera le plus en pratique sera le cardinal de l'ensemble correspond (ici 1).

Définition 4. Il existe un vocabulaire précis pour certains événements particuliers :

- L'événement Ω est appelé **événement certain**. C'est de fait un événement qui se produira toujours.
- L'événement vide est appelé **événement impossible** et n'est jamais réalisé.
- Un événement est élémentaire s'il est constitué d'un seul élément de Ω .
- Deux événements sont **incompatibles** si leur intersection est vide (autrement dit, ils ne peuvent pas être réalisés simultanément).
- Un **système complet d'événements** est un ensemble d'événements deux à deux incompatibles, et dont la réunion vaut Ω (autrement dit, une partition de Ω).

Exemples :

- L'événement A cité ci-dessus est un événement élémentaire.
- L'événement B : « On tire deux piques, deux cœurs et deux trèfles » est un événement impossible.
- Les événements C : « On tire au moins un pique et au moins un carreau » et D : « On tire cinq cartes de la même couleur » sont incompatibles.
- Les événements E_0 : « On ne tire pas d'As » ; E_1 : « On tire exactement un As » ; ... ; E_4 : « On tire quatre As » forment un système complet d'événements.

1.3 Tribus

Définition 5. Soit Ω un univers et $\mathcal{T} \subset \mathcal{P}(\Omega)$ un ensemble de sous-ensembles de Ω . On dit que \mathcal{T} est une tribu sur Ω si

- $\Omega \in \mathcal{T}$
- \mathcal{T} est stable par passage au complémentaire (si $A \in \mathcal{T}$ alors $\bar{A} \in \mathcal{T}$) et par union dénombrable (si $\forall i \in \mathbb{N}, A_i \in \mathcal{T}$, alors $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{T}$).

Remarque 2. Une tribu est également stable par intersection dénombrable en utilisant la stabilité par passage au complémentaire et les lois de De Morgan.

Remarque 3. La tribu représentera en fait l'ensemble des événements pour lesquels on saura calculer une probabilité. Comme il est facile de calculer la probabilité d'un complémentaire et d'une union (voire plus loin dans ce cours les formules correspondantes), les conditions imposées sont en fait assez naturelles. Quand Ω est fini, on prendra toujours $\mathcal{T} = \mathcal{P}(\Omega)$ car on sait calculer des probabilités pour tous les sous-ensembles. C'est quand Ω est infini que le concept de tribu devient essentiel.

Exemple : Dans le cas où $\Omega = \mathbb{R}$ (ou \mathbb{R}^+), une tribu très fréquemment utilisée est celle des *boréliens*, qui est constituée de toutes les unions dénombrables d'intervalles. En particulier, elle contient tous les intervalles.

1.4 Lois de probabilité

Définition 6. On appelle **espace probabilisable** un couple (Ω, \mathcal{T}) , où Ω est un univers et \mathcal{T} une tribu sur Ω .

Définition 7. Une **probabilité** sur un espace probabilisable (Ω, \mathcal{T}) est une application $P : \mathcal{T} \rightarrow [0; 1]$ vérifiant :

- $P(\Omega) = 1$
- si $(A_i)_{i \in \mathbb{N}}$ est une suite d'événements deux à deux incompatibles de \mathcal{T} , $P(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} P(A_i)$

Remarque 4. La deuxième propriété, appelée σ -additivité, est souvent utilisée pour une union finie plutôt que dénombrable.

Définition 8. Un **espace probabilisé** est un triplet (Ω, \mathcal{T}, P) , où P est une probabilité sur l'espace probabilisable (Ω, \mathcal{T}) .

Remarque 5. On peut très bien avoir envie de mettre plusieurs probabilités sur un même espace probabilisable. Prenons l'exemple simple d'un lancer de dé. La probabilité « naturelle » consiste à décréter que $P(1) = P(2) = \dots = P(6) = \frac{1}{6}$ (il suffit de définir les probabilités des événements élémentaires puisqu'on obtient les probabilités des autres événements par σ -additivité). Mais ce n'est pas la seule ! Par exemple, $P(1) = \frac{1}{21}$; $P(2) = \frac{2}{21}$; \dots ; $P(6) = \frac{6}{21}$ définit également une probabilité (la seule chose à vérifier est que $P(\Omega)$ vaut 1, ce qu'on obtient en faisant la somme des probabilités des événements élémentaires).

2 Propriétés

2.1 Généralités

Proposition 1. Si P est une loi de probabilité, on a toujours $P(\emptyset) = 0$.

Démonstration. L'événement vide étant incompatible avec lui-même (c'est bien le seul à vérifier cette curieuse propriété !), il doit vérifier $P(\emptyset) + P(\emptyset) = P(\emptyset \cup \emptyset) = P(\emptyset)$, donc $P(\emptyset) = 0$. \square

Proposition 2. Pour tout événement $A \in \mathcal{T}$, on a $P(\bar{A}) = 1 - P(A)$.

Démonstration. Les événements A et \bar{A} sont incompatibles, donc on a $P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(\Omega) = 1$, ce qui donne bien le résultat voulu. \square

Proposition 3. Si $A \subset B \in \mathcal{T}^2$, on a $P(A) \leq P(B)$.

Démonstration. On peut écrire, de façon similaire à la démonstration précédente, $P(A) + P(B \setminus A) = P(B)$. Or $P(B \setminus A) \geq 0$ (une probabilité est toujours comprise entre 0 et 1), donc on a bien $P(B) \geq P(A)$. \square

Proposition 4. Soient $(A, B) \in \mathcal{T}^2$, alors $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Plus généralement,

la formule de Poincaré est valable : $P(\bigcup_{i=1}^n A_i) = \sum_{k=1}^n ((-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \dots \cap A_{i_k}))$.

Démonstration. On a par σ -additivité $P(A) = P(A \cap B) + P(A \setminus B)$; $P(B) = P(A \cap B) + P(B \setminus A)$ et $P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B)$. L'égalité souhaitée en découle. Comme dans le cas des ensembles, on se gardera de faire une démonstration complète de la formule de Poincaré (qui se démontre d'ailleurs de la même façon que dans le cas des ensembles). \square

Proposition 5. Soient $(A_i)_{i \in \mathbb{N}}$ un système complet d'événements, alors $\sum_{i \in \mathbb{N}} P(A_i) = 1$ et $\forall B \in \mathcal{T}$,

$$P(B) = \sum_{i \in \mathbb{N}} P(B \cap A_i).$$

Démonstration. Cela découle en fait immédiatement de la définition. Comme les A_i sont par définition incompatibles, $P(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} P(A_i)$. Or, la réunion des A_i vaut Ω (par définition également), donc sa probabilité vaut 1. La formule pour $P(B)$ est similaire, il suffit de remarquer que les ensembles $B \cap A_i$ sont disjoints et que leur réunion est égale à B (en fait, ils forment un système complet d'événements pour B). \square

Remarque 6. Une formule similaire est valable dans le cas d'un système complet d'événements fini.

2.2 Probabilités sur un univers fini

Remarque 7. Répétons une remarque importante déjà faite dans un précédent paragraphe : dans le cas où Ω est fini, pour déterminer une probabilité, il suffit de connaître la probabilité de chaque événement élémentaire, avec la seule condition que la somme de ces probabilités soit égale à 1. En effet, tout événement est une union finie disjointe de tels événements élémentaires.

Définition 9. Il y a **équiprobabilité** sur l'espace Ω si tous les événements élémentaires ont la même probabilité.

Proposition 6. Dans le cas de l'équiprobabilité, on a simplement, $\forall A \in \mathcal{P}(\Omega)$, $P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$.

Démonstration. Posons $n = |\Omega|$. Les événements élémentaires formant un système complet d'événements, la somme de leur probabilités vaut 1. Or, cette somme est constituée de n nombres égaux (par définition de l'équiprobabilité), donc chacune de ces probabilités vaut $\frac{1}{n}$. Ensuite, un événement quelconque est union disjointe des événements élémentaires qui le composent, sa probabilité vaut donc k fois $\frac{1}{n}$, où k est le nombre d'éléments dans cet événement. \square

Exemple Si nous reprenons notre exemple favori, la probabilité d'avoir une quinte flush (cinq cartes qui se suivent dans la même couleur) est très faible : 4 quintes possibles dans chaque couleur, soit 16 cas favorables, à diviser par $\binom{32}{5}$. Cela donne une proba d'environ 0.000 08.

3 Probabilités conditionnelles

Le principe des probabilités conditionnelles est, si on y réfléchit bien, assez simple, et surtout utilisé sans forcément qu'on s'en rende compte dans nombre d'exercices. Comme son nom l'indique, la notion désigne une probabilité soumise à une condition. Prenons un exemple simple : on lance deux dés et on regarde leur somme (vous devez commencer à avoir l'habitude). On a vu dans le chapitre précédent que la probabilité d'obtenir 5 valait $\frac{1}{9}$. Supposons qu'on ait maintenant l'information supplémentaire : on sait que le premier dé est tombé sur la face 2. Ça change tout ! Pour obtenir un total de 10, il faut maintenant (et il suffit) que le deuxième dé tombe sur 3, soit une chance sur 6. On dit que la probabilité d'obtenir 5 sachant que le premier dé a donné 2 vaut $\frac{1}{6}$ (naturellement, il sera plus commode de noter ceci à l'aide d'événements).

3.1 Notations

Définition 10. Soit (Ω, \mathcal{T}, P) un espace probabilisé et A, B deux événements tels que $P(A) \neq 0$. La **probabilité conditionnelle** de B sachant A est $P_A(B) = \frac{P(A \cap B)}{P(A)}$.

Remarque 8. Si on veut être savant, on peut dire que l'application de \mathcal{T} dans $[0, 1]$ définie par $B \mapsto \frac{P(A \cap B)}{P(A)}$ définit une nouvelle probabilité sur \mathcal{T} , appelée probabilité sachant A et notée P_A .

Remarque 9. On rencontre souvent la notation alternative $P(B|A)$ pour la probabilité conditionnelle, mais nous ne l'utiliserons pas dans ce cours.

Exemple : Cela correspond bien à l'idée intuitive. Sachant que A est réalisé, on se place dans un nouvel univers constitué des événements vérifiant A , et dans ce nouvel univers, B est réalisé pour tous les événements de $A \cap B$, ce qui conduit à la formule de la définition. Si on reprend l'exemple introductif, en notant A : « Le premier dé donne 2 » et B : « Le total vaut 5 », on a $P(A) = \frac{1}{6}$ et $P(A \cap B) = \frac{1}{36}$ (il n'y a qu'un cas qui marche, celui où on obtient 2 et 5), donc la probabilité conditionnelle vaut bien $\frac{1}{6}$.

Remarque 10. La probabilité conditionnelle étant une loi de probabilité, elle a les mêmes propriétés que n'importe quelle probabilité, en particulier $P_A(\bar{B}) = 1 - P_A(B)$, ou $P_A(B \cup C) = P_A(B) + P_A(C) - P_A(B \cap C)$.

3.2 Théorèmes

Théorème 1. Formule des probabilités composées. Soient (A_1, A_2, \dots, A_n) des événements tels que $P(\bigcap_{i=1}^{n-1} A_i) \neq 0$, alors $P(\bigcap_{i=1}^n A_i) = P(A_1) \times P_{A_1}(A_2) \times P_{A_1 \cap A_2}(A_3) \times \dots \times P_{A_1 \cap A_2 \cap \dots \cap A_{n-1}}(A_n)$.

Remarque 11. La condition demandée sert simplement à assurer que toutes les probabilités conditionnelles sont bien définies : en effet, si $P(\bigcap_{i=1}^{n-1} A_i) \neq 0$, on a a fortiori $P(A_1) \neq 0$; $P(A_1 \cap A_2) \neq 0$, etc.

Remarque 12. Si on écrit la formule dans le cas où il n'y a que deux événements, on obtient simplement $P(A_1 \cap A_2) = P(A_1) \times P_{A_1}(A_2)$, ce qui est simplement la définition d'une probabilité conditionnelle.

Démonstration. On va procéder par récurrence. D'après la première remarque, toutes les probabilités conditionnelles sont bien définies, et d'après la seconde la formule est vérifiée pour $n = 2$. Supposons-la vraie au rang n , on a alors $P(\bigcap_{i=1}^{n+1} A_i) = P(\bigcap_{i=1}^n A_i) \times P_{A_1 \cap A_2 \cap \dots \cap A_n}(A_{n+1})$ (on a simplement utilisé la formule dans le cas d'une intersection de deux événements), et via l'hypothèse de récurrence, on obtient immédiatement le résultat. \square

Cette formule est utilisée presque systématiquement dans les cas où on a des tirages chronologiques, et correspond simplement à la formalisation de la représentation intuitive sous forme d'arbre. Deux exemples parmi tant d'autres :

Exemple : l'exemple bateau, genre de calcul qu'on fait en permanence sans invoquer les probabilités composées. Dans une urne se trouvent 4 boules blanches et 3 noires. On tire successivement trois boules. La probabilité d'obtenir une noire, puis deux blanches vaut $\frac{3}{7} \times \frac{4}{6} \times \frac{3}{5} = \frac{6}{35}$.

Exemple : On dispose de trois enveloppes contenant respectivement 3 pièces d'un euro ; 5 pièces d'un euro et 3 de deux euros ; 4 d'un euro. On choisit une enveloppe au hasard, puis une pièce au hasard dans l'enveloppe. La probabilité d'obtenir une pièce de 2 euros vaut $\frac{1}{3} \times \frac{3}{8} = \frac{1}{8}$.

Théorème 2. Formule des probabilités totales. Si les événements A_i forment un système complet d'événements et si $\forall i, P(A_i) \neq 0$, alors pour tout événement B , $P(B) = \sum P(A_i) \times P_{A_i}(B)$ (j'ai volontairement omis de préciser dans quel ensemble se baladait l'indice i pour éviter d'avoir plusieurs cas à traiter).

Remarque 13. Dans le cas d'un système complet de deux événements, on obtient la forme plus simple suivante : $P(B) = P(A) \times P_A(B) + P(\bar{A}) \times P_{\bar{A}}(B)$.

Démonstration. Dans le cas particulier, on a vu plus haut que $P(B) = P(A \cap B) + P(\bar{A} \cap B)$. Or, on sait que $P(A \cap B) = P(A) \times P_A(B)$, et de même pour la deuxième moitié. Le cas général se fait exactement de la même manière. \square

Cette formule est très utile dans les cas où l'expérience aléatoire se déroule en deux (ou plus) étapes avec à la fin de la première étape une partition des possibilités en plusieurs cas disjoints (c'est-à-dire encore une fois quand on fait une représentation sous forme d'arbre).

Exemple : Dans une urne se trouvent 4 boules noires et 6 blanches. On tire successivement trois boules dans l'urne, sachant qu'après chaque tirage on remet la boule tirée, mais qu'on en ajoute une autre de la même couleur. La probabilité d'obtenir une boule noire au premier tirage vaut $\frac{2}{5}$, et celle d'obtenir une boule noire au deuxième tirage vaut $\frac{4}{10} \times \frac{5}{11} + \frac{6}{10} \times \frac{4}{11} = \frac{2}{5}$. Au troisième tirage, elle vaut aussi $\frac{2}{5}$. Pas si évident que ça à justifier sans calcul.

Exemple : Un type de problème très classique en probabilités et faisant intervenir les probabilités totales est la chaîne de Markov. Il s'agit d'une situation qui évolue au cours du temps, et pour laquelle la situation à un instant donné ne dépend que de la situation à l'instant précédent (mais de manière aléatoire, tout de même). Par exemple, Homer Simpson mange tous les matins au petit déjeuner soit un beignet, soit un croissant. Au jour numéroté 0, il a mangé un beignet. S'il mange un beignet au jour n , il mangera un croissant au jour $n + 1$ avec probabilité $\frac{1}{2}$, et à nouveau un beignet avec probabilité $\frac{1}{2}$. Par contre, s'il mange un croissant, il passera au beignet le lendemain avec probabilité $\frac{2}{3}$ et reprend un croissant avec probabilité $\frac{1}{3}$. On cherche à déterminer en fonction de n la probabilité qu'Homer mange un beignet au jour n . Notons donc a_n la probabilité qu'il mange un beignet au jour n et b_n celle qu'il mange un croissant. On a par hypothèse $a_0 = 1$ et $b_0 = 0$, et ensuite, en utilisant la formule des probabilités totales, les relations de récurrence $a_{n+1} = \frac{1}{2}a_n + \frac{2}{3}b_n$ et $b_{n+1} = \frac{1}{2}a_n + \frac{1}{3}b_n$ (cette deuxième relation ne sert d'ailleurs à rien pour répondre à notre question). Comme on sait par ailleurs que $a_n + b_n = 1$ (Homer ne saute jamais son petit-déjeuner), on obtient $a_{n+1} = \frac{1}{2}a_n + \frac{2}{3}(1 - a_n) = \frac{2}{3} - \frac{1}{6}a_n$. La suite (a_n) est donc une suite arithmético-géométrique, d'équation de point fixe $x = \frac{2}{3} - \frac{1}{6}x$, ce qui donne $x = \frac{2}{3} \times \frac{6}{7} = \frac{4}{7}$. Posons donc $v_n = a_n - \frac{4}{7}$, on a alors $v_{n+1} = a_{n+1} - \frac{4}{7} = -\frac{1}{6}a_n + \frac{2}{3} - \frac{4}{7} = -\frac{1}{6}a_n + \frac{2}{21} = -\frac{1}{6}\left(a_n - \frac{4}{7}\right) = -\frac{1}{6}v_n$. La suite (v_n) est donc géométrique de raison $-\frac{1}{6}$ et de premier terme $v_0 = a_0 - \frac{4}{7} = \frac{3}{7}$. On en déduit que $a_n = v_n + \frac{4}{7} = \frac{4}{7} + \frac{3}{7}\left(-\frac{1}{6}\right)^n$. Notons au passage que, quand n tend vers $+\infty$, la limite de la suite a_n vaut $\frac{4}{7}$. Autrement dit, à long terme, Homer mange des beignets en moyenne quatre jours par semaine.

Théorème 3. Formule de Bayes. Soient A et B deux événements de probabilités non nulles, alors
$$P_B(A) = \frac{P(A) \times P_A(B)}{P(B)}.$$

Démonstration. Il n'y a presque rien à faire, on sait que $P(A \cap B) = P(A) \times P_A(B) = P(B) \times P_B(A)$, la formule en découle immédiatement. \square

Remarque 14. On peut également donner une version de cette formule avec plus de deux événements, mais cela ne présente guère d'intérêt pratique.

Cette formule, qui n'apporte en apparence pas grand chose par rapport aux précédentes, est en fait très utile dans la mesure où elle permet de « remonter le temps » lorsqu'on a une expérience faisant apparaître des choix chronologiques.

Exemple : On peut reprendre n'importe quel exemple classique, mais en essayant de faire les choses dans le sens inhabituel. On tire deux dés successivement, le total obtenu est 9. Quelle est la probabilité que le premier dé soit tombé sur 4? Notons A : « Le premier dé tombe sur 5 » et B : « Le total obtenu est 9 ». On a $P(A) = \frac{1}{6}$, $P(B) = \frac{1}{9}$ et $P_A(B) = \frac{1}{6}$, donc $P_B(A) = \frac{1}{4}$.

4 Indépendance d'événements

Encore une notion relativement intuitive a priori, mais qui nécessite une définition mathématique précise. On a envie de dire que deux événements sont indépendants si la réalisation de l'un n'influence pas celle de l'autre. Par exemple, en reprenant comme souvent notre lancer de deux dés, l'événement « Le premier dé tombe sur un chiffre pair » devrait logiquement être indépendant de l'événement « Le deuxième dé tombe sur 1 ou 2 ».

Définition 11. Deux événements A et B sont **indépendants** si $P(A \cap B) = P(A) \times P(B)$.

Proposition 7. Deux événements de probabilité non nulle sont indépendants si et seulement si $P(B) = P_A(B)$.

Démonstration. Rappelons que sous ces hypothèses $P(A \cap B) = P(A) \times P_A(B)$. En identifiant les deux formules, on obtient tout de suite le résultat. \square

Remarque 15. Dans le cas où l'un des deux événements a une probabilité nulle, les événements sont de toute façon nécessairement indépendants (même si c'est absurde!).

Exemple On tire une carte dans un jeu de 32 cartes. Les événements « obtenir un Roi » et « obtenir un pique » sont indépendants.

Remarquez que, dans le but de prouver l'indépendance de deux événements, cette formulation n'est pas vraiment plus simple à utiliser que l'autre. Passons au cas de plusieurs événements.

Définition 12. Des événements (A_1, A_2, \dots, A_n) sont dits **mutuellement indépendants** si $\forall I \subset \{1, 2, \dots, n\}$, on a $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$.

Remarque 16. Attention, des événements mutuellement indépendants sont indépendants deux à deux mais la réciproque n'est pas vraie! La condition est beaucoup plus forte que ça. Par exemple, pour trois événements, on doit avoir $P(A_1 \cap A_2) = P(A_1) \times P(A_2)$; $P(A_1 \cap A_3) = P(A_1) \times P(A_3)$; $P(A_2 \cap A_3) = P(A_2) \times P(A_3)$ mais aussi $P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3)$. Les conditions à vérifier deviennent rapidement affreuses quand on augmente le nombre d'événements.

Proposition 8. Si A et B sont indépendants, A et \bar{B} le sont aussi.

Démonstration. C'est un petit calcul : $P(\bar{A} \cap B) = P(B \setminus A) = P(B) - P(A \cap B) = P(B) - P(A) \times P(B) = P(B) \times (1 - P(A)) = P(B) \times P(\bar{A})$. \square

Remarque 17. Ce résultat se généralise à plus de deux événements : si (A_1, A_2, \dots, A_n) sont mutuellement indépendants, on peut remplacer une partie des A_i par leur complémentaire et conserver l'indépendance mutuelle.

Exemple : On fait une série de n lancers de pièces. Les événements A_k : « On obtient Pile au k ème lancer » sont mutuellement indépendants.

Exemple : On lance deux dés (si, si, je vous jure, encore une fois), et on considère les événements A : « Le premier dé donne un résultat pair », B : « Le deuxième dé donne un résultat pair » et C : « Les deux dés donnent des résultats de même parité ». Je vous laisse vérifier que $P(A) = P(B) = P(C) = \frac{1}{2}$ et $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, donc les événements sont deux à deux indépendants. Pourtant, $P(A \cap B \cap C) = \frac{1}{4}$, donc les trois événements ne sont pas mutuellement indépendants.