Chapitre 42

Espace probabilisé fini

42.1 Le langage de l'aléatoire

§1 Notion d'événement et d'univers

Pour modéliser la notion de résultat aléatoire d'une expérience, le plus simple, pour un mathématicien (qui n'a pas à économiser sur le matériel), est de construire des mondes différents, chacun donnant un des résultats possibles de l'expérience.

Prenons un exemple simple : on lance un dé à six faces. La main jette le dé qui décrit une élégante parabole avant de frapper la surface d'une table. À ce moment là, l'histoire bifurque, et de nombreux futurs sont possibles. Si l'on se contente d'une description partielle du lancer de dé, on a six «futurs possibles», chacun d'eux étant caractérisé par la face visible une fois le dé immobile. Une description plus complète des rebonds successifs du dé entraînera au contraire une multiplication des futurs possibles. Les amateurs de science-fiction peuvent songer à tous les ω comme à des mondes parallèles.

Définition 1 Soit $\mathcal{E}$ une expérience aléatoire, nous appellerons **univers** associé à $\mathcal{E}$, l'ensemble noté traditionnellement Ω de tous les résultats possibles de $\mathcal{E}$.
Le **résultat** de l'expérience est l'un des éléments ω de l'univers Ω. Pour des raisons historique, ils portent de nombreux noms : épreuve, aléa, résultat de l'expérience, réalisation du hasard, état du monde, résultat d'une expérience...

On prendra bien garde au fait que, entre une expérience réelle et sa transcription mathématique (choix de Ω puis description d'une probabilité sur Ω), il y a une étape de **modélisation**. Cette étape est la seule partie non mathématique de la théorie des probabilités ; elle est évidemment sujette à caution et procède de choix.

Exemple 2

Tirage d'un dé
Considérons un personnage qui, à un instant t_0, lance un dé à six faces. On peut, par exemple, décider que les éléments de Ω sont les descriptions complètes du monde réel (avec tous ses atomes, leur évolution dans le temps). De (très) nombreux éléments de Ω correspondent à un même tirage de dé. Bien entendu, une telle description est bien trop complexe et, pour tout dire, inutile.
De manière beaucoup plus économique, on peut «oublier» la plus grosse partie du monde, n'en conserver que les propriétés essentielles et choisir une modélisation plus simple, dans laquelle on décrète que $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$. C'est évidement ce point de vue qui sera adopté dans la suite.

Exemple 3

Arrêt d'un programme
On considère un programme acceptant en entrée une série de données de taille arbitraire, on note T le temps (compté en nombre de calculs élémentaires effectués par l'ordinateur) avant l'arrêt du programme. On pose alors $\Omega = \mathbb{N}^\star \cup \{ \infty \}$.

Exemple 4

Succession de tirages à pile ou face
Si l'on accepte l'idée d'une succession infinie de tirages au sort, on prend comme univers l'ensemble $\Omega = \{ P, F \}^{\mathbb{N}^\star}$ des applications de $\mathbb{N}^\star$ dans $\{ P, F \}$, c'est-à-dire des suites $\left(r_n\right)_{n\geq 1}$ à valeurs dans $\{ P, F \}$. Dans une suite $\omega = \left(r_n\right)_{n\geq 1}$ donnée, la valeur de r_k est le résultat du k-ième tirage.

Soit $\mathcal{E}$ une expérience aléatoire et Ω un l'univers qui lui est associé. On dit qu'un événement est lié à l'expérience $\mathcal{E}$ si pour tout résultat $\omega \in \Omega$, on sait dire si cet événement a eu lieu ou non.[1]

Nous conviendrons de représenter et même d'identifier un tel événement par l'ensemble de $\omega \in \Omega$ pour lesquels il a eu lieu. Un événement lié à $\mathcal{E}$ sera donc identifié à une partie de Ω.

Définition 5

- Les **événements** de l'expérience aléatoire sont les parties de l'ensemble Ω.
- Les **événements élémentaires** de l'expérience aléatoire sont les singletons de l'ensemble Ω.

Si ω est élément de A, on dit que **l'événement A se réalise dans l'épreuve ω**. Très couramment, on ne fait pas mention de ω, et l'on dit simplement que «l'événement A est réalisé».

Exemple 6

Revenons au cas du tirage d'un dé ; l'univers choisi est $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$. Voici quelques événements:

- «le résultat est pair» qui est l'événement $\{ 2, 4, 6 \}$. On dira alors que, lors d'une expérience de lancer de dé (c'est-à-dire lors du choix d'une réalisation particulière ω du hasard), «le résultat est pair» pour signifier «$\omega \in \{ 2, 4, 6 \}$».
- «le résultat est impair» qui est l'événement $\{ 1, 3, 5 \}$.
- «le résultat est supérieur ou égal à 5» qui est l'événement $\{ 5, 6 \}$.

[1] On n'a pas défini le mot événement !

- L'événement certain qui est l'événement Ω tout entier. Il apparaît de façon naturelle (par exemple «le résultat est inférieur à 42»).
- L'événement impossible qui est l'événement $\emptyset$: aucun résultat d'expérience n'est fait partie. Lui aussi apparaît de façon naturelle (par exemple : «le résultat est négatif»).

Toute les parties de Ω (il y en a $2^6 = 64$) sont des événements.

§2 Langage ensembliste

La définition d'un événement comme partie de Ω nous permet de transcrire, en langage ensembliste, certaines locutions intuitives. Ainsi, la phrase : *l'événement A est réalisé dans l'épreuve ω* se traduit simplement par «ω est élément de A».

De même la phrase «la réalisation de A implique la réalisation de B» se traduit par

$$\forall \omega \in \Omega, \omega \in A \implies \omega \in B$$

c'est-à-dire $A \subset B$.

Notation	Langage ensembliste	Langage probabiliste
ω ($\omega \in \Omega$)	élément	résultat, épreuve, éventualité,...
A ($A \subset \Omega$)	partie de Ω (cas fini)	événement
$\{ \omega \}$ ($\omega \in \Omega$)	singleton	événement élémentaire
$\omega \in A$	ω est élément de A	A est réalisé (dans le résultat ω)
$A \subset B$	A inclus dans B	l'événement A implique l'événement B
$A \cup B$	réunion	réalisation de A *ou* de B
$A \cap B$	intersection	réalisation de A *et* de B
A^c ou $\Omega \setminus A$	complémentaire	non-A (événement contraire)
$A \setminus B$	A privé de B	réalisation de A mais pas de B
$\emptyset$	partie vide	événement impossible
Ω		événement certain
$A \cap B = \emptyset$	parties disjointes	A et B sont des événements incompatibles

Rappel On rappelle quelques propriétés des opérations sur les événements

$$A \cup (B \cup C) = (A \cup B) \cap (A \cup C) \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup A^c = \Omega \qquad A \cap A^c = \emptyset$$
$$(A^c)^c = A \qquad \Omega^c = \emptyset$$
$$(A \cup B)^c = A^c \cap B^c \qquad (A \cap B)^c = A^c \cup B^c$$

ces deux dernières égalités portant le nom de **règles de de Morgan**.

42.2 Espace probabilisé fini

§1 Espace probabilisable

Définition 7

Soit Ω un ensemble non vide. Une **tribu** $\mathcal{T}$ sur Ω est une partie de $\mathcal{P}(\Omega)$ telle que

- $\Omega \in \mathcal{T}$,
- pour toute partie de $A \in \mathcal{T}$, le complémentaire A^c appartient à $\mathcal{T}$,
- pour tout suite $\left(A_n\right)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$ à valeurs dans $\mathcal{T}$,

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}.$$

Le couple $(\Omega, \mathcal{T})$ est alors appelé **espace probabilisable**. Les éléments de $\mathcal{T}$ sont appelés **événements**.

Exemple 8

L'ensemble $\mathcal{P}(\Omega)$ constitue une tribu sur Ω.

§2 Espace probabilisé fini

Dans le cadre du programme de première année, l'univers Ω est fini. On choisit donc en général comme tribu $\mathcal{T} = \mathcal{P}(\Omega)$.

Définition 9

Une **probabilité** (ou **mesure de probabilité**) sur un *univers fini* Ω est une application de P définie sur l'ensemble des événements $\mathcal{T}$ ($\mathcal{T}$ comme **tribu**) vérifiant :

- P est à valeurs dans $\mathbb{R}_+$;
- l'univers Ω est de probabilité unité : $P(\Omega) = 1$;
- P est **additive**, c'est-à-dire que si A et B sont des événements disjoints, alors la probabilité associée à leur union est la somme de leurs probabilités :

$$P(A \underset{\text{disj.}}{\cup} B) = P(A) + P(B).$$

Le triplet $(\Omega, \mathcal{T}, P)$ est un **espace probabilisé fini**.

Dans ce cas, le réel $P(A)$ est appelé **probabilité de l'événement A**.

Remarque

Cette liste d'axiome est suffisante lorsque Ω est un ensemble fini ; nous devrons la compléter par un axiome supplémentaire lorsque Ω est infini.

Une **probabilité** (ou **mesure de probabilité**) sur l'espace probabilisable $(\Omega, \mathcal{T})$ est une application de P définie sur l'ensemble des événements $\mathcal{T}$ vérifiant :

- P est à valeurs dans $\mathbb{R}_+$;
- l'univers Ω est de probabilité unité : $P(\Omega) = 1$;

- P est **σ-additive**, c'est-à-dire que pour toute suite $\left(A_n\right)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$ d'événements deux à deux disjoints, la série $\sum_{n\in\mathbb{N}} P(A_n)$ converge et
$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

Le triplet $(\Omega, \mathcal{T}, P)$ est un **espace probabilisé**.

Remarque Un événement A est dit **négligeable** ou s'il est de probabilité nulle ; $P(A) = 0$. Il est dit **presque sûr** s'il est de probabilité 1. Une propriété est dite vraie **presque sûrement** si l'ensemble des éléments de Ω vérifiant cette propriété est un événement de probabilité 1.

§3 Propriétés des mesures de probabilités

Proposition 10 **Propriétés élémentaires**

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé fini. Soient A et B deux parties de Ω.

1. $P(\emptyset) = 0$;
2. *P est à valeurs dans* $[0, 1]$
3. *Si $A \subset B$, $P(A) \leq P(B)$ (croissance de P);*
4. *Si $A \subset B$, $P(B \setminus A) = P(B) - P(A)$;*
5. $P(A^c) = 1 - P(A)$;
6. $P(B \setminus A) = P(B) - P(A \cap B)$.

Corollaire 11

1. *Si $A_1, \ldots, A_n$ sont des événement deux à deux disjoints, alors*
$$P(A_1 \underset{\text{disj.}}{\cup} A_2 \underset{\text{disj.}}{\cup} \ldots \underset{\text{disj.}}{\cup} A_n) = P(A_1) + P(A_2) + \cdots + P(A_n). \qquad \text{(additivité)}$$
2. *Si $A_1, \ldots, A_n$ sont des événements quelconques, alors*
$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n). \qquad \text{(sous-additivité)}$$

Théorème 12 *Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé fini. Soient A et B deux parties de Ω. Alors*
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

§4 Probabilités sur un ensemble fini

Munir un univers fini $\Omega = \{ \omega_1, \dots, \omega_n \}$ d'une probabilité revient à associer à chaque élément ω_k, un réel $p_k \geq 0$, avec pour seule contrainte $p_1 + p_2 + \cdots + p_n = 1$. Insistons sur le fait que l'on *choisit* effectivement ces nombres, et qu'on ne les calcule pas. Le choix peut être plus ou moins pertinent, c'est-à-dire que la modélisation est plus ou moins réussie.

Définition 13 Une **distribution de probabilités** sur un ensemble E est une famille d'éléments de $\mathbb{R}_+$ indexée par E et de somme 1.

Le théorème suivant montre qu'une probabilité P sur Ω est déterminée par la distribution de probabilités $(P(\{ \omega \}))_{\omega \in \Omega}$.

Théorème 14 *Soit $\Omega = \{ \omega_1, \omega_2, \dots, \omega_n \}$ et $p_1, p_2, \dots, p_n$ des nombres réels.*
Il existe une probabilité P sur $(\Omega, \mathcal{T})$ telle que pour tout $k \in [\![1, n]\!]$, $P(\{ w_k \}) = p_k$ si, et seulement si

- $\forall i \in [\![1, n]\!], p_i \geq 0$,
- $\sum_{i=1}^{n} p_i = 1$.

P est alors unique, et on a pour tout événement A,

$$P(A) = \sum_{\omega \in A} P(\{ \omega \}) = \sum_{\substack{k=1\dots n \\ \omega_k \in A}} p_k.$$

Exemple 15 Dans le cas d'un lancer de dé, on peut choisir $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ et poser $p_k = 1/6$ pour $k = 1, \dots, 6$; cette situation modélise un dé équilibré. Si l'on se rend compte que, pour un dé donné, les prédictions du modèle ne se réalisent pas, on peut modifier les valeurs des coefficients p_k. Pour déterminer expérimentalement chaque coefficient, une méthode est de lancer un très grand nombre de fois le dé et de calculer des fréquences expérimentales. Des résultats théoriques permettent de donner une estimation du nombre de lancers à effectuer pour obtenir des valeurs suffisamment précises des p_k.

§5 Systèmes complets d'événements

Définition 16 Une famille $(A_1, \dots, A_n)$ de parties non vides de Ω est appelée **système complet d'événements** si elle forme une partition de Ω, c'est-à-dire que Ω est la réunion disjointe des A_i :

$$\Omega = A_1 \underset{\text{disj.}}{\cup} \dots \underset{\text{disj.}}{\cup} A_n.$$

Exemple 17 Un événement A et son contraire A^c, tous deux non vides, constituent un système complet d'événement (A, A^c) de Ω.

Proposition 18 **Formule des probabilités totales V1**

Soit $(A_1, \dots, A_n)$ est un système complet d'événements de Ω et B un événement de Ω, alors

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n).$$

En particulier si l'on a un événement A et son contraire A^c, alors

$$P(B) = P(B \cap A) + P(B \cap A^c).$$

Exemple 19 On prend au hasard une chaussette dans la commode de Jack. Il y a une chance sur neuf pour que ce soit une chaussette trouée rouge et deux chances sur sept pour que ce soit une chaussette rouge sans trou. On note A l'événement « la chaussette est rouge » et B l'événement « la chaussette est trouée ». L'énoncé nous donne les probabilités $P(A \cap B) = \frac{1}{9}$ et $P(A \cap B^c) = \frac{2}{7}$. Nous avons ainsi la probabilité de prendre une chaussette rouge

$$P(A) = P(A \cap B) + P(A \cap B^c) = \frac{1}{9} + \frac{2}{7} = \frac{25}{63}.$$

§6 Équiprobabilité

Définition 20 Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé fini.
On dit qu'il y a **équiprobabilité**, lorsque les probabilités de tous les événements élémentaires sont égales. On dit aussi que P est la **probabilité uniforme** sur $(\Omega, \mathcal{T})$.

Théorème 21 *Soit $(\Omega, \mathcal{T})$ un espace probabilisable fini.*
L'hypothèse d'équiprobabilité définit sur cet espace une probabilité P unique, donnée par

$$\forall A \subset \Omega, P(A) = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)}.$$

*On dit aussi que P est **la probabilité uniforme** sur $(\Omega, \mathcal{T})$.*

Remarque Lorsque l'on effectue des tirages **au hasard**, on sous-entend qu'il y a équiprobabilité.

Dans ce cas, le calcul des probabilités se ramène alors à deux dénombrements.

Exemple 22 On lance deux dés ordinaires bien équilibrés. Il y a équiprobabilité. Ω est l'ensemble des couples d'entiers de un à six. Il y a donc $6^2 = 36$ éventualités. L'événement A «un as et un six» a deux éventualités : $(1, 6)$ et $(6, 1)$. Il a pour probabilité

$$P(A) = \frac{\mathrm{card}\, A}{\mathrm{card}\, \Omega} = \frac{2}{36} = \frac{1}{18}.$$

Il est parfois plus simple de considérer l'événement contraire.

Exemple 23

On lance trois pièces bien équilibrées. Il y a équiprobabilité. Ω est l'ensemble des 3-listes (ou triplets) d'éléments de $\{ P, F \}$ (pile, face). Il y a donc $2^3 = 8$ éventualités.
L'événement A «au moins une fois pile» a pour contraire $A^c = \{ (F, F, F) \}$. Sa probabilité est donc

$$P(A) = 1 - P(A^c) = 1 - \frac{\operatorname{card} A^c}{\operatorname{card} \Omega} = 1 - \frac{1}{8} = \frac{7}{8}.$$

42.3 Conditionnement

Un petit paradoxe Considérons l'expérience consistant à lancer deux pièces (discernables et équilibrées), représentée par l'espace probabilisé fini

$$\Omega = \{ (F, F), (F, P), (P, F), (P, P) \},$$

muni de l'équiprobabilité. Soient l'événement A : «face est sorti» et l'événement B: «pile est sorti». Ils sont représentés par les parties

$$A = \{ (F, F), (F, P), (P, F) \}, \quad \text{et} \quad B = \{ (F, P), (P, F), (P, P) \}.$$

En comptant le nombre d'élément de ces deux parties, on obtient facilement les probabilités de ces événements: $P(A) = P(B) = 3/4$.

Supposons maintenant que l'on ait vu qu'une des pièces a donné face. Dans ce cas, il semble clair que B est réalisé si, et seulement si l'autre pièce donne pile, ce qui a une chance sur deux d'être réalisé. L'information «A est réalisé» a donc modifié la probabilité que l'on va attribuer à l'événement B, et l'on obtient alors

$$P\,(B \text{ sachant que } A \text{ est réalisé}) = 1/2. \qquad \text{(Faux)}$$

Examinons maintenant la situation d'un point de vue ensembliste. Si l'une des pièces a donné F, l'événement A est réalisé; on est donc dans le cas où l'éventualité ω qui correspond à ce lancer est telle que $\omega \in A$. Dans ce cas, il est clair que B est réalisé si, et seulement si $\omega \in A \cap B$. Et, dans notre exemple on obtient

$$P\,(B \text{ sachant que } A \text{ est réalisé}) = \frac{\operatorname{card}(A \cap B)}{\operatorname{card}(A)} = 2/3.$$

La discordance paradoxale entre les deux résultats que l'on vient d'obtenir nous incite à la prudence…

§1 Probabilités conditionnelles

Définition 24

Soit B un événement de probabilité non nulle : $P(B) > 0$. Pour tout événement A, on définit la **probabilité conditionnelle de A sachant B** comme le réel, noté $P_B(A)$ ou $P(A|B)$, valant

$$P_B(A) = P(A|B) \overset{\text{def}}{=} \frac{P(A \cap B)}{P(B)}.$$

Exemple 25

Dans le bureau de Gaston, à la pause-café, cinq femmes sur sept et trois hommes sur cinq prennent leur café sans sucre. On prend une de ces douze personnes au hasard. On note F l'événement «la personne est une femme» et S l'événement «la personne prend son café sans sucre». On a $P(S) = \frac{8}{12}$ et $P(F \cap S) = \frac{5}{12}$, par conséquent la probabilité pour que ce soit une femme sachant qu'elle prend son café sans sucre est

$$P(F|S) = \frac{P(F \cap S)}{P(S)} = \frac{5}{8}.$$

Le point essentiel de cette définition est que

Théorème 26

L'application

$$\begin{array}{rccc} P_B : & \mathcal{T} & \rightarrow & [0,1] \\ & A & \mapsto & P(A|B) \end{array}$$

est une probabilité sur $(\Omega, \mathcal{T})$ *; on l'appelle* ***probabilité conditionnelle sachant B*** *ou encore* ***probabilité conditionnée par B****.*

Exemple 27

Jack possède une commode Louis XV en noyer à trois tiroirs. Sa chaussette rose a une chance sur deux de se trouver dans ce meuble (indifféremment dans l'un des tiroirs) et une chance sur deux de se trouver à l'extérieur. On a ouvert les deux premiers tiroirs: elle n'y est pas! Sachant cela, quelle est la probabilité qu'elle soit dans le troisième tiroir?
Il semble naturel de prendre l'espace probabilisé (Ω, P)

extérieur	tiroir 1	tiroir 2	tiroir 3
1/2	1/6	1/6	1/6

On conditionne par l'événement B:«on a ouvert les deux premiers tiroirs, elle n'y est pas», qui est représenté par la partie $\{$ extérieur, tiroir 3 $\}$. On considère l'événement A représenté par la partie $\{$ tiroir 3 $\}$ et l'on a

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{\text{ tiroir } 3\})}{P(\{\text{ extérieur, tiroir } 3\})} = \frac{1/6}{1/2 + 1/6} = 1/4.$$

Conditionner par B revient à considérer l'espace probabilisé (Ω, P_B):

extérieur	tiroir 1	tiroir 2	tiroir 3
3/4	0	0	1/4

Remarque

La définition de la probabilité conditionnelle par la formule

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

est très simple, mais il est important de bien l'interpréter. Il faut bien voir dans cette formule que $A|B$ *n'est pas un événement* et que $P(A|B)$ n'est autre qu'une notation pour $P_B(A)$. L'interprétation de $P_B(A)$ n'est pas la même que celle de $P(A \cap B)$, et *ces deux probabilités ne sont pas égales* (sauf si $P(B) = 1$).

Proposition 28

Soit B un événement de probabilité non nulle : $P(B) > 0$. *Alors*

1. $P(B|B) = 1$.

2. *Si* $B \subset A$*,* $P(A|B) = 1$.

3. $P\left(\emptyset|B\right) = 0$.

4. $P(A|B) + P(A^c|B) = 1$.

De manière générale, on a

$$P(A|B) + P(A|B^c) \neq 1.$$

Cette somme pouvant prendre une valeur quelconque entre 0 et 2 (Exemple : je joue au loto; je note B l'événement «je perd ma mise» et A l'événement «il fait beau»; la somme précédente est dramatiquement proche de 2).

Remarque **Sur les notations $P_B(A)$ et $P(A|B)$**

Dans les ouvrages anglo-saxons, les probabilités conditionnelles sont presque toujours notées «$P(A|B)$»; l'usage s'est répandu en France de les noter «$P_B(A)$». À ceci, on trouve deux avantages. Le premier est de pouvoir parler de la mesure de probabilité elle-même, notée P_B. Le second est d'ordre pédagogique: il permet de rappeler qu'il s'agit de travailler avec une autre probabilité que celle qui est initialement définie, et de ne pas oublier qu'*il n'existe pas d'événement «$(A|B)$»*.

§2 Formule des probabilités composées

Théorème 29 *Soit* B *un événement de probabilité non nulle :* $P(B) > 0$*. Alors*

$$P(A \cap B) = P(B) \cdot P(A|B).$$

Le programme autorise la convention $P(A|B)P(B) = 0$ lorsque $P(B) = 0$.

Exemple 30 Dans son panier, Jeanne a trois pommes véreuses et quatre pommes sans ver. Elle prend deux pommes une par une, sans les remettre. Quelle est la probabilité de prendre la première sans ver et la seconde véreuse?

Notons A l'événement «la première n'a pas de ver» et B «la deuxième est véreuse». Par équiprobabilité du premier tirage, $P(A) = \frac{4}{7}$. Si la première n'est pas véreuse, il reste trois pommes véreuses et trois pommes sans ver pour le deuxième tirage: $P(B|A) = \frac{3}{6} = \frac{1}{2}$. Donc

$$P(A \cap B) = P(A)P(B|A) = \frac{2}{7}.$$

Théorème 31 **Formule des probabilités composées**

Soit $n \geq 2$ *un entier et soient* $A_1, A_2, \dots, A_n$ *des événements ; on suppose que* $A = A_1 \cap A_2 \cap \cdots \cap A_{n-1}$ *est de probabilité non nulle. Alors*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

Démonstration. Une simple récurrence. ∎

Exemple 32 On tire successivement 4 cartes d'un jeu de 52 cartes. Quelle est la probabilité de tirer les 4 as?

Notons A_i ($i = 1, \dots, 4$) l'événement «la i-ème carte tirée est un as». On demande de calculer $P(A_1 \cap A_2 \cap A_3 \cap A_4)$. Les probabilité immédiatement accessible sont, en-dehors de $P(A_1) = \frac{4}{52}$, les probabilités conditionnelles

$$P(A_2|A_1) = \frac{3}{51} \qquad P(A_3|A_1 \cap A_2) = \frac{2}{50} \qquad P(A_4|A_1 \cap A_2 \cap A_3) = \frac{1}{49}.$$

On a donc

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3 \cap A_4) &= P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot P(A_4|A_1 \cap A_2 \cap A_3) \\ &= \frac{4 \cdot 3 \cdot 2 \cdot 1}{52 \cdot 51 \cdot 50 \cdot 49} \approx 0.00000369. \end{aligned}$$

On peut retrouver ce résultat par dénombrement, mais on observera que nous ne nous sommes pas préoccupés dans cet exemple de la construction d'un espace Ω dont les A_i pourraient être considérés comme des sous-ensembles.

§3 Formule des probabilité totales

Le résultat suivant, bien qu'élémentaire, est d'une grande importance dans le calcul des probabilités.

Théorème 33 **Formule des probabilités totales**
Soient $(A_1, \dots, A_n)$ une famille formant un système complet d'événements. Soit B un événement. On suppose $P(A_i) > 0$ pour tout i. Alors

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

Corollaire 34 *Si $P(A) \neq 0$ et $P(A^c) \neq 0$,*

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c).$$

Exemple 35 On considère l'énoncé suivant.
La probabilité qu'il pleuve un matin donné est $p = 1/5$. Lorsqu'il pleut, la probabilité que je prenne mon parapluie est $9/10$. Lorsqu'il ne pleut pas, la probabilité que je le prenne est $2/10$. Quelle est la probabilité que je prenne mon parapluie un matin donné?

Démonstration. La première étape dans la résolution d'un tel problème est de formaliser les données. On note A l'événement «il pleut», et A^c l'événement «il ne pleut pas»; l'univers Ω est bien sûr égal à l'union disjointe $A \cup A^c$. Notons B l'événement «je prends mon parapluie». La phrase «Lorsqu'il pleut, la probabilité que je prenne mon parapluie est $9/10$» doit s'interpréter comme une *probabilité conditionnelle*, c'est-à-dire comme $P(B|A)$ («sachant qu'il pleut»), et non comme la probabilité d'une conjonction $P(B \cap A)$ («il pleut *et* je prends mon parapluie»). On a ainsi

$$P(B|A) = \frac{9}{10} \qquad \text{et} \qquad P(B|A^c) = \frac{2}{10}.$$

La formule des probabilités totales donne alors

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c) = \frac{1}{5}\cdot\frac{9}{10} + \frac{4}{5}\cdot\frac{2}{10} = \frac{17}{50}.$$

∎

Exemple 36

Dans le championnat de Liverpool d'échecs, Paul, John et Georges vont disputer la sélection finale dont le vainqueur sera opposé au champion en titre Robert. Paul a une chance sur trois d'être sélectionné, John une sur deux et, par conséquent, Georges une sur six. Contre Paul, Robert a deux chances sur trois de gagner, contre John trois sur cinq et contre Georges, trois sur quatre. Quelle est la probabilité pour Robert de remporter le titre?
Les événements P=«Paul est sélectionné», J=«John est sélectionné» et G=«Georges est sélectionné» forment un système complet d'événements. Notons R l'événement «Robert remporte le titre». Les hypothèses

$$\begin{aligned} P(P) &= \frac{1}{3} & P(J) &= \frac{1}{2} & P(G) &= \frac{1}{6} \\ P(R|P) &= \frac{2}{3} & P(R|J) &= \frac{3}{5} & P(R|G) &= \frac{3}{4}. \end{aligned}$$

nous permettent de calculer

$$\begin{aligned} P(R) = P(P)\cdot P(R|P) + P(J)\cdot P(R|J) + P(G)\cdot P(R|G) \\ = \frac{1}{3}\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{3}{5} + \frac{1}{6}\cdot\frac{3}{4} = \frac{233}{360}. \end{aligned}$$

§4 Formule de Bayes

Théorème 37

Théorème de Bayes v.1
Soient A et B deux événements de probabilité non nulle. On a alors

$$P(A|B) = \frac{P(A)\cdot P(B|A)}{P(B)}.$$

Exemple 38

«Tu finiras sur l'échafaud!» s'exclame une grand-tante excédée à l'adresse de son petit-neveu qui s'est gavé de bonbons et s'obstine à le nier. «On commence par mentir, et on devient criminel!», ajoute-t-elle, car elle a de la culture et a lu dans son hebdomadaire favori que 90% des criminels ont commencé par mentir effrontément durant leur jeunesse (alors que seuls 20% des gens honnêtes ont menti étant jeunes).
Sachant qu'il n'y a que 1% de criminel dans la population, peut-on calculer la probabilité qu'a le petit enfant de devenir criminel et rassurer ainsi la brave grand-tante?
Notons C l'événement «être criminel» et M l'événement «mentir durant son enfance». Les données chiffrées connues sont

$$P(C) = 0.01 \quad P(M|C) = 0.9 \quad P(M|C^c) = 0.2.$$

Le théorème de Bayes nous permet d'écrire

$$P(C|M) = \frac{P(C) \cdot P(M|C)}{P(M)},$$

mais il nous manque la valeur de $P(M)$. On peut l'obtenir par la formule des probabilités totales en utilisant le système complet d'événements (C, C^c):

$$P(M) = P(M|C)P(C) + P(M|C^c)P(C^c),$$

ce qui donne

$$P(C|M) = \frac{P(C) \cdot P(M|C)}{P(M|C)P(C) + P(M|C^c)P(C^c)}.$$

L'application numérique donne

$$P(C|M) = \frac{0.01 \cdot 0.9}{0.01 \cdot 0.9 + 0.2 \cdot 0.99} \approx 0.043 \approx 4.3\%.$$

La grand-tante peut s'endormir tranquille.

Théorème 39

Théorème de Bayes

Soient $(A_1, \dots A_n)$ un système complet d'événements et B un événement. On suppose que $P(A_i) > 0$ pour tout i et $P(B) > 0$; alors

$$\forall p \in [\![1, n]\!], P(A_p|B) = \frac{P(A_p) \cdot P(B|A_p)}{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}.$$

42.4 Indépendance stochastique

§1 Indépendance de deux événements

Définition 40

Deux événements A et B sont **indépendants** (ou **stochastiquement indépendants**) s'ils vérifient

$$P(A \cap B) = P(A) \cdot P(B).$$

Remarque

L'indépendance de deux événements est une notion qui dépend de la probabilité P dont on a muni l'espace $(\Omega, \mathcal{T})$, au contraire de l'incompatibilité, qui est une propriété ne dépendant que de Ω.

Proposition 41

Soient A et B deux événements.

1. *Si A est un événement de probabilité non nulle, alors A et B sont indépendants si, et seulement si $P(B|A) = P(B)$.*
2. *Les évenement $\emptyset$ et Ω sont indépendants de tout autre événement.*
3. *A est indépendant de lui-même si, et seulement si il est de probabilité 0 ou 1.*

4. *Si A et B sont indépendants, alors A et B^c el sont, ainsi que A^c et B, et que A^c et B^c.*

5. *Si A est de probabilité nulle, alors A est indépendant de B.*

§2 Indépendance mutuelle

Si A, B, C deux à deux indépendants, a-t-on $P(A \cap B \cap C) = P(A)P(B)P(C)$?

Exemple 42

Considérons un lancer de deux dés et les événements

- A : la somme des deux dés est paire.
- B : le premier dé est pair.
- C : le deuxième dé est pair.

Ces événements sont bien deux à deux indépendants, puisque

$$P(A) = P(B) = P(C) = \frac{1}{2} \quad \text{et} \quad P(A \cap B) = P(B \cap C) = P(C \cap A) = \frac{1}{4}.$$

En revanche, $P(A \cap B \cap C) = P(B \cap C) = \frac{1}{4}$ alors que $P(A)P(B)P(C) = \frac{1}{8}$.

On doit introduire une définition plus restrictive pour caractériser l'indépendance de plusieurs événements.

Définition 43

Des événements $A_1, A_2, \dots, A_n$ sont dit (**mutuellement**) **indépendants** si, pour tout entier $p \le n$ et pour tout p-uplet d'indices $(i_1, \dots, i_p)$ deux à deux distincts, on a

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_p}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_p}).$$

Plus généralement, si I est un ensemble d'indices et soit $(A_i)_{i \in I}$ une famille d'événements. On dit que ces événements sont **mutuellement indépendants** si pour toute partie *finie* $J \subset I$ non vide, on a

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Remarque

- Lorsque $n = 2$, on retrouve évidemment la définition précédente. Lorsque $n = 3$, les événements A, B, C sont mutuellement indépendants si, et seulement si l'on a les quatre relations

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \qquad & P(A \cap B \cap C) = P(A)P(B)P(C) \\ P(A \cap C) &= P(A)P(C) & \\ P(B \cap C) &= P(B)P(C) & \end{aligned}$$

- L'indépendance mutuelle implique l'indépendance deux à deux.

Proposition 44 *Soit $(A_1, \dots, A_n)$ une famille quelconque d'événement mutuellement indépendants. On considère la famille d'événement $(B_1, \dots, B_n)$ telle que*

$$B_i = A_i \quad \text{ou} \quad B_i = A_i^c.$$

Alors $(B_1, \dots, B_n)$ est une famille d'événements mutuellement indépendants.