27

Espaces probabilisés finis et variables aléatoires

Nous développons dans ce chapitre des outils pour le calcul des probabilités.
Bien que vous ayez déjà développé une excellente compréhension de ce qu'est[1] une probabilité, la principale difficulté est de comprendre le formalisme utilisé, qui s'inscrit toujours dans le cadre[2] de la théorie des ensembles.
Il n'y aura donc aucun hasard dans la suite, nous ne ferons que le **modéliser**, et tous les dés que nous évoquerons, toutes les urnes dans lesquelles nous tirerons des boules ne seront que purement imaginaires.

[1] Ou au moins de ce que doit être.

[2] Totalement déterministe.

S'il serait extrêmement réducteur de cantonner la théorie des probabilités aux lancers de dés, tirages de boules dans des urnes et autre jeux de hasard plus ou moins tordus, ces situations faciles à comprendre présentent tout de même un véritable intérêt pédagogique de par leur simplicité.
Je ne me priverai donc pas d'utiliser de tels exemples, même s'il faudra garder à l'esprit qu'il ne s'agit là que de *toy models*.
La physique regorge de situations intéressantes où il faut être capable de calculer des probabilités, ce qui nécessite déjà une certaine familiarité avec le concept de probabilité.

Enfin, conformément aux programmes en vigueur, nous ne manipulerons cette année que des expériences qui ne possèdent qu'un nombre fini d'issues, et il vous faudra attendre l'année prochaine pour aller plus loin.

27.1 Espaces probabilisés finis

27.1.1 Univers, événement

Définition 27.1 – On appelle **univers fini** un ensemble fini Ω non vide.
Les éléments de Ω sont appelés **événements élémentaires**.

Par exemple, un lancer de pièce sera modélisé par $\{P, F\}$ ou $\{0, 1\}$, un lancer de dé sera modélisé par $\Omega = [\![1, 6]\!]$, le lancer de deux dés sera modélisé par $\Omega = [\![1, 6]\!]^2$ et le tirage d'une main de 5 cartes d'un jeu de 52 cartes sera modélisé par $\Omega = \mathscr{P}_5\left(\left\{\boxed{2\spadesuit}, \boxed{2\diamondsuit}, \dots, \boxed{A\clubsuit}, \boxed{A\heartsuit}\right\}\right)$, l'ensemble des parties à 5 éléments de l'ensemble des 32 cartes.

Notez bien que l'univers qui va modéliser l'expérience est le même que la pièce soit équilibrée ou non, nous n'avons pas encore parlé de probabilité !
La plupart du temps, l'univers sur lequel on travaille n'est pas complètement clair, ou tout du moins, il n'est pas nécessaire de l'expliciter pour pouvoir faire des calculs.
L'idée est que l'univers est l'ensemble de toutes les issues possibles à une expérience aléatoire, et plus l'expérience est complexe, plus l'univers risque de l'être.
Par exemple si une expérience consiste à lancer un dé à 6 faces, puis si on a obtenu la face portant le numéro k, à placer $2k$ boules blanches et $3k$ boules rouges dans une urne, puis en tirer avec remise k boules, on n'a pas vraiment envie d'expliciter Ω, l'ensemble de toutes les issues possibles à l'expérience[3]. Ce qui ne nous empêchera pas de calculer des probabilités avec cette expérience !

La restriction sur la finitude du cardinal de Ω n'est pas une vraie limitation, même si le cas des univers infinis requiert un peu plus de précautions[4]. Nous nous limiterons donc cette

[3] L'une de ces issues serait par exemple «tomber sur la face numéro 4 puis obtenir 3 boules blanches puis une boule rouge lors de nos tirages».

[4] Car il nous amène à manipuler des sommes infinies, que nous n'avons pas encore définies et qui sont plus délicates à manipuler que les sommes finies.

année aux univers fini afin de prendre le temps de nous habituer au formalisme.
Cette contrainte fait qu'on ne pourra par exemple pas s'intéresser cette année à une expérience du type «lancer une pièce jusqu'à obtenir un premier PILE», puisqu'il y a une infinité d'issues possibles à cette expérience (obtenir PILE au premier lancer, obtenir PILE au second lancer,... ,obtenir PILE au 10^6-ème lancer, etc).
Vous manipulerez des univers infinis en spé, mais pas encore dans le cadre le plus général[5].

[5] Et par exemple vous ne rencontrerez pas en prépa les variables aléatoire dites à densité (loi exponentielle, loi normale...).

En particulier, nous n'étudierons pas cette année de situations où il y a une infinité d'issues possibles, par exemple le fait de lancer une pièce jusqu'à obtention du premier pile.

Dans toute la suite du chapitre, on considère un univers fini Ω fixé.

Définition 27.2 – On appelle **événement** toute partie de l'univers Ω, c'est-à-dire tout élément de $\mathscr{P}(\Omega)$.

Par abus de langage, on appellera encore événement élémentaire tout singleton $\{\omega\} \in \mathscr{P}(\Omega)$, sans vraiment distinguer ω et $\{\omega\}$.
Ainsi, un événement est un ensemble d'événements élémentaires, ou pour le dire autrement, un ensemble d'issues possibles de l'expérience considérée.

Exemples 27.3

- Lorsqu'on lance un dé, l'événement «obtenir un nombre pair» est $\{2, 4, 6\}$.
- Lorsqu'on lance deux dés, l'événement «la somme des dés vaut 6» est $\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$.
- Lorsqu'on tire au main au poker, l'événement «avoir un full» est un ensemble à 3744 éléments[6] : $\left\{\left\{\boxed{A\spadesuit}, \boxed{A\diamondsuit}, \boxed{A\heartsuit}, \boxed{J\spadesuit}, \boxed{J\clubsuit}\right\}, \dots, \left\{\boxed{7\clubsuit}, \boxed{7\diamondsuit}, \boxed{7\heartsuit}, \boxed{Q\spadesuit}, \boxed{Q\heartsuit}\right\}\right\}$

[6] Voir le cours de dénombrement.

Définition 27.4 – Le complémentaire dans Ω d'un événement A est appelé le **contraire** de A, et noté $\overline{A}$.
Si A et B sont deux événements, on dit que $A \cup B$ est la disjonction de A et B, plus souvent appelé «A **ou** B», et $A \cap B$ est la conjonction de A et B, plus souvent appelé «A **et** B».
On dit que A et B sont **incompatibles** si $A \cap B = \emptyset$.

Classique
L'exemple le plus typique d'événements incompatibles est celui d'un événement A et de son contraire $\overline{A}$.

Il faut bien comprendre que $A \cup B$ est l'événement «l'un au moins des deux événements A et B est réalisé».
Et de même, $A \cap B$ est l'événement «A et B sont réalisés simultanément».
Pour rester sur des exemples simples, lançons un dé, et soit $A = \{2, 4, 6\}$ l'événement «obtenir un numéro pair» et $B = \{4, 5, 6\}$ l'événement «obtenir un numéro supérieur ou égal à 4».
Alors $A \cap B = \{4, 6\}$ est bien l'événement «obtenir un numéro pair supérieur ou égal à 4», alors que $A \cup B = \{2, 4, 5, 6\}$ contient bien toutes les issues de l'expérience qui réalisent soit A soit B (soit les deux en même temps).
Des événements incompatibles sont donc des événements qui ne peuvent pas se réaliser en même temps.

Exemples 27.5

Considérons une urne, qui contient a boules blanches et b boules noires. On tire n boules de cette urne, avec remise entre les tirages.
Pour tout $i \in [\![1, n]\!]$, notons B_i (resp. N_i) l'événement «la $i^{\text{ème}}$ boule tirée est blanche (resp. noire)».
Alors :

- $B_i = \overline{N_i}$

Univers
Un univers possible pour cette expérience est $\{B, N\}^n$, l'ensemble des suites de n éléments à valeurs dans $\{B, N\}$ (pour Noir et Blanc).
B_i est alors l'ensemble de tous les n-uplets dont le $i^{\text{ème}}$ élément est un B.
Vous remarquerez que le nombre de boules n'entre pas en jeu tant qu'on n'a pas parlé de probabilité de chacun des événements.

- $\displaystyle\bigcup_{i=1}^{n} N_i$ est l'événement «obtenir au moins une boule noire»
- $\displaystyle\bigcap_{i=1}^{n} B_i$ est l'événement «n'obtenir que des blanches», contraire du précédent
- $\displaystyle\bigcup_{i=1}^{n-1} (B_i \cap B_{i+1})$ est l'événement «obtenir deux boules blanches consécutives».
- les événements

$$\overline{\left(\bigcup_{i=1}^{n-1}(B_i \cap B_{i+1})\right) \cup \left(\bigcup_{i=1}^{n-1}(N_i \cap N_{i+1})\right)} \text{ et } \left(\bigcap_{\substack{i=1\\ i \text{ pair}}}^{n} N_i \cap \bigcap_{\substack{j=1\\ j \text{ impair}}}^{n} B_j\right) \cup \left(\bigcap_{\substack{i=1\\ i \text{ pair}}}^{n} B_i \cap \bigcap_{\substack{j=1\\ j \text{ impair}}}^{n} N_j\right)$$

sont égaux : ce sont tous les deux l'événement «deux boules consécutives sont de couleurs différentes».

Définition 27.6 – On appelle **système complet d'événements** tout ensemble $\{A_1, \dots, A_n\}$ d'événements, deux à deux incompatibles, et tels que $\displaystyle\bigcup_{i=1}^{n} A_i = \Omega$.
Autrement dit, un système complet d'événements est un recouvrement disjoint de Ω.

Abréviation
Il est courant d'abréger *système complet d'événements* en *s.c.e.*
Bien entendu cette abréviation n'est pas tolérée dans une copie... du moins c'est le discours officiel.

Exemples 27.7

- Si A est un événement, alors $\{A, \overline{A}\}$ est un système complet d'événements.
- $\{\{\omega\}, \omega \in \Omega\}$ est un système complet d'événements.
- Lors d'un lancer de dé, si on note A l'événement «obtenir un numéro pair» et B l'événement «obtenir un numéro impair», alors $\{A, B\}$ est un système complet d'événements.
- Lors du lancer de deux dés, si on note A_i l'événement «la somme des deux dés vaut i», alors $\{A_i, 2 \leqslant i \leqslant 12\}$ est un système complet d'événements.

27.1.2 Probabilités

Nous aimerions désormais, pour chaque événement, pouvoir parler de sa probabilité, qu'on veut donc être un réel entre $[0, 1]$.

Définition 27.8 – Soit Ω un univers fini. On appelle **probabilité** sur Ω toute application $\mathbf{P} : \mathscr{P}(\Omega) \to [0, 1]$ telle que :

- $\mathbf{P}(\Omega) = 1$
- $\forall (A, B) \in \mathscr{P}(\Omega)^2,\ A \cap B = \emptyset \Rightarrow \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ *(additivité).*

Pour $A \in \mathscr{P}(\Omega)$, le réel $\mathbf{P}(A)$ est appelé probabilité de l'événement A.
On appelle alors **espace probabilisé fini**[7] un couple $(\Omega, \mathbf{P})$ où Ω est un univers fini et $\mathbf{P}$ une probabilité sur Ω.

⚠ Attention !
Ne pas confondre la probabilité $\mathbf{P}$, qui est une application, de la probabilité $\mathbf{P}(A)$ d'un événement A, qui est un réel.

[7] En abrégé «espace probabilisé» puisque cette année nous ne parlerons que du cas fini.

Remarque. Demander que $\mathbf{P}$ soit à valeurs dans $\mathbf{R}_+$ plutôt que $[0, 1]$ suffirait, car alors pour tout $A \in \mathscr{P}(\Omega)$, $1 = \mathbf{P}(\Omega) = \mathbf{P}(A) + \underbrace{\mathbf{P}(\overline{A})}_{\geqslant 0}$ et donc $\mathbf{P}(A) \leqslant 1$.

Quand on lance un dé, c'est au moment de choisir la probabilité que l'on met sur $[\![1, 6]\!]$ que l'on décide si notre dé est ou non équilibré.
Par exemple, dans le cas[8] où le dé tombe absolument toujours sur 6, la probabilité qu'on

[8] Un peu extrême et complètement inintéressant je vous le concède.

utilise sur $\mathscr{P}([\![1,6]\!])$ est définie par

$$\mathbf{P}(A) = \begin{cases} 0 & \text{si } 6 \notin A \\ 1 & \text{si } 6 \in A \end{cases} = \mathbb{1}_A(6).$$

Je ne prouve pas qu'il s'agit bien là d'une probabilité, mais nous y reviendrons.

L'exemple le plus naturel de probabilité, qui est celui qui nous vient en premier à l'esprit est celui de probabilité uniforme, qui correspond à ce qu'on appelle une situation d'équiprobabilité, où l'on considère que toutes les issues élémentaires de l'expérience ont la même probabilité de se produire.

Proposition 27.9 : *Soit Ω un univers fini. Alors il existe une unique probabilité $\mathbf{P}$ sur Ω telle que $\omega \mapsto \mathbf{P}(\{\omega\})$ soit constante*[9]*.*
Cette probabilité, appelée ***probabilité uniforme sur*** *Ω est donnée par*

$$\forall A \in \mathscr{P}(\Omega),\ \mathbf{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}.$$

[9] C'est-à-dire telle que tous les événements élémentaires (= toutes les issues de l'expérience) soient de même probabilité.

Démonstration. Supposons qu'une telle probabilité $\mathbf{P}$ existe, et soit $\lambda \in [0,1]$ la valeur commune à tous les $\mathbf{P}(\{\omega\})$, $\omega \in \Omega$.
Alors

$$\mathbf{P}(\Omega) = \mathbf{P}\left(\bigcup_{\omega\in\Omega}\{\omega\}\right) = \sum_{\omega\in\Omega}\mathbf{P}(\{\omega\}) = \sum_{\omega\in\Omega}\lambda = \lambda\mathrm{Card}(\Omega).$$

Évident ?
Le fait de passer de la proba de l'union à la somme des probas doit vous sembler logique, la vraie justification se trouve à la proposition suivante.

Puisque $\mathbf{P}(\Omega) = 1$, on a donc $\lambda = \dfrac{1}{\mathrm{Card}(\Omega)}$.
Et en particulier, pour tout événement $A \in \mathscr{P}(\Omega)$, il vient

$$\mathbf{P}(A) = \mathbf{P}\left(\bigcup_{\omega\in A}\{\omega\}\right) = \sum_{\omega\in A}\mathbf{P}(\{\omega\}) = \mathrm{Card}(A)\frac{1}{\mathrm{Card}(\Omega)}.$$

À présent, définissons une application $\mathbf{P}$ sur $\mathscr{P}(\Omega)$ par $\mathbf{P}(A) = \dfrac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$.
Alors il est évident que $\mathbf{P}$ est à valeurs dans $[0,1]$ (puisqu'un cardinal est positif, et que le cardinal d'une partie A de Ω est inférieur ou égal à $\mathrm{Card}(\Omega)$), et on a bien $P(\Omega) = \dfrac{\mathrm{Card}(\Omega)}{\mathrm{Card}(\Omega)} = 1$.
Enfin, si A et B sont deux événements incompatibles,

$$\mathbf{P}(A\cup B) = \frac{\mathrm{Card}(A\cup B)}{\mathrm{Card}(\Omega)} = \frac{\mathrm{Card}(A)+\mathrm{Card}(B)}{\mathrm{Card}(\Omega)}.$$

Donc $\mathbf{P}$ est bien une probabilité sur Ω. □

Exemples 27.10

► Au poker, si l'on considère que toutes les mains sont équiprobables[10], la probabilité d'avoir un carré est $\dfrac{624}{2\,598\,960} \simeq 0.000\,24$.
► Si je choisis au hasard une personne en cours de maths, la probabilité qu'elle soit née un 27 juillet est $\frac{3}{47}$.

[10] Ce qui semble légitime si vous ne vivez pas dans un album de Lucky Luke.

Théorème 27.11 (Calculs avec des probabilités) : *Soit* $\mathbf{P}$ *une probabilité sur* Ω. *Alors :*

1. $\mathbf{P}(\emptyset) = 0$
2. $\forall A \in \mathscr{P}(\Omega)$, $\mathbf{P}\left(\overline{A}\right) = 1 - \mathbf{P}(A)$
3. $\forall (A, B) \in \mathscr{P}(\Omega)^2$, $A \subset B \Rightarrow \mathbf{P}(A) \leqslant \mathbf{P}(B)$ *(croissance de la probabilité)*
4. $\forall (A, B) \in \mathscr{P}(\Omega)^2$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
5. $\forall (A_1, \ldots, A_n) \in \mathscr{P}(\Omega)^n$, $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mathbf{P}(A_i)$ *(sous-additivité).*
   Si de plus les A_i *sont deux à deux disjoints, alors on a égalité :*
   $$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}(A_i).$$

Autrement dit
$\mathbf{P}$ est une application croissante de l'ensemble ordonné $(\mathscr{P}(\Omega), \subset)$ dans $([0,1], \leqslant)$.

Démonstration. 1. On a $\Omega \cap \emptyset = \emptyset$, et donc par additivité,
$$\mathbf{P}(\Omega) = \mathbf{P}(\Omega \cup \emptyset) = \mathbf{P}(\Omega) + \mathbf{P}(\emptyset) \text{ donc } \mathbf{P}(\emptyset) = 0.$$

2. On a $1 = \mathbf{P}(\Omega) = \mathbf{P}(A \cup \overline{A}) = \mathbf{P}(A) + \mathbf{P}(\overline{A})$ car A et $\overline{A}$ sont incompatibles.
   Et donc $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.
3. Si $A \subset B$, alors $B = A \cup (B \setminus A)$, et l'union est disjointe.
   Donc $\mathbf{P}(B) = \mathbf{P}(A) + \underbrace{\mathbf{P}(B \setminus A)}_{\geqslant 0} \geqslant \mathbf{P}(A)$.
4. Puisque $A \cup B = B \cup (A \setminus B)$, il vient $\mathbf{P}(A \cup B) = \mathbf{P}(B) + \mathbf{P}(A \setminus B)$.
   Mais $A = (A \setminus B) \cup (A \cap B)$, et encore une fois il s'agit d'une union disjointe, donc
   $$\mathbf{P}(A) = \mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) \text{ si bien que } \mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$$
   et donc $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
5. Par récurrence sur n, le nombre d'événements prouvons $\mathscr{P}(n)$:
   «$\forall (A_1, \ldots, A_n) \in \mathscr{P}(\Omega)^n$, $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mathbf{P}(A_i)$ et si les A_i sont deux à deux incompatibles, alors $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}(A_i)$.»
   Pour $n = 2$, on a bien $\mathbf{P}(A_1 \cup A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cap A_2) \leqslant \mathbf{P}(A_1) + \mathbf{P}(A_2)$.
   Et le cas d'incompatibilité découle directement de la définition de probabilité.
   Supposons donc $\mathscr{P}(n)$ vraie, et soient $A_1, \ldots, A_{n+1} \in \mathscr{P}(\Omega)$. Alors
   $$\mathbf{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbf{P}\left(A_1 \cup \bigcup_{i=2}^{n+1} A_i\right) \leqslant \mathbf{P}(A_1) + \mathbf{P}\left(\bigcup_{i=2}^{n+1} A_i\right) \leqslant \mathbf{P}(A_1) + \sum_{i=2}^{n+1} \mathbf{P}(A_i) \leqslant \sum_{i=1}^{n+1} \mathbf{P}(A_i).$$
   Dans le cas où les A_i sont deux à deux incompatibles, alors A_1 est incompatible avec $\bigcup_{i=2}^{n+1} A_i$, et donc les deux inégalités ci-dessus sont des égalités[11].
   Ainsi, $\mathscr{P}(n+1)$ est vraie, et par le principe de récurrence, la propriété est vraie pour tout n.

□

[11] La première par définition d'une probabilité, la seconde par hypothèse de récurrence.

Remarques. ► La signification de l'inclusion n'est pas complètement évidente au premier abord. Si $A \subset B$, cela signifie que tous les événements élémentaires qui composent A sont dans B.
Autrement dit, que **dès que** A **est réalisé, alors** B **l'est automatiquement.** C'est vraiment ainsi qu'il faut comprendre l'inclusion.
Dès lors, la croissance de la probabilité se comprend bien : B doit donc être plus souvent réalisé que A et donc avoir une probabilité supérieure à celle de A.

▸ Le point 4. concernant la probabilité d'une union ne doit pas du tout vous surprendre dans le cas d'équiprobabilité, puisque c'est alors une conséquence directe de

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B).$$

▸ Sur le même principe que pour les cardinaux, on peut généraliser le point 3 :

$$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C)$$

et il existe une formule du crible[12] qui est exactement la même que celle vue en TD pour les cardinaux, mais en remplaçant les cardinaux par des probabilités.

[12] Hors-programme.

27.1.3 Construction de probabilités

Une conséquence facile du dernier point de la proposition qui précède est que si $\mathbf{P}$ est une probabilité, alors $\displaystyle\sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) = 1$.

Nous nous intéressons ici à une réciproque à ce résultat :

Proposition 27.12 : *Soit $\Omega = \{\omega_1, \dots, \omega_n\}$ un ensemble fini, et soient $p_1, \dots, p_n$ des réels positifs tels que* $\displaystyle\sum_{i=1}^{n} p_i = 1$.
Alors il existe une unique probabilité $\mathbf{P}$ sur Ω telle que $\forall i \in [\![1, n]\!]$, $\mathbf{P}(\{\omega_i\}) = p_i$.

Remarque
Si les p_i sont des nombres **positifs** de somme 1, alors tous sont compris entre 0 et 1.

Démonstration. Pour l'unicité, commençons par noter que pour tout $A \in \mathscr{P}(\Omega)$, $\mathbf{P}(A) = \displaystyle\sum_{x \in A} \mathbf{P}(\{x\})$, et donc si deux probabilités coïncident sur tous les événements élémentaires, alors elles sont égales.

Considérons à présent l'application $\mathbf{P}$ définie sur $\mathscr{P}(\Omega)$ par $\mathbf{P}(A) = \displaystyle\sum_{i=1}^{n} p_i \mathbb{1}_A(\omega_i)$.

Autrement dit
Les p_i qui apparaissent dans la somme sont ceux pour lequels $\omega_i \in A$.

Alors $\mathbf{P}(\Omega) = \displaystyle\sum_{i=1}^{n} p_i = 1$. Si A et B sont deux événements incompatibles, nous savons que $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \underbrace{\mathbb{1}_{A \cap B}}_{=0}$, et donc

$$\mathbf{P}(A \cup B) = \sum_{i=1}^{n} p_i \mathbb{1}_{A \cup B}(\omega_i) = \sum_{i=1}^{n} p_i \mathbb{1}_A(\omega_i) + \sum_{i=1}^{n} p_i \mathbb{1}_B(\omega_i) = \mathbf{P}(A) + \mathbf{P}(B).$$

Donc $\mathbf{P}$ est bien une probabilité, qui a clairement les propriétés requises. □

Ce que légitime cette proposition, c'est que se donner une probabilité sur Ω revient à se donner les probabilités de tous les événements élémentaires, sous l'hypothèse très raisonnable que ces probabilités soient bien positives et de somme 1.
Par exemple, il n'existe qu'une probabilité sur $[\![1, 6]\!]$ qui permet de simuler un lancer d'un dé qui tomberait sur $2, 4$ ou 6 avec probabilité $\frac{1}{6}$, sur 1 ou 3 avec probabilité $\frac{1}{10}$ et sur 5 avec probabilité $\dfrac{3}{10}$.

27.2 PROBABILITÉS CONDITIONNELLES

27.2.1 Définition

Définition 27.13 – Soient A et B deux événements d'un espace probabilisé fini $(\Omega, \mathbf{P})$, tels que $\mathbf{P}(A) \neq 0$.
On appelle **probabilité conditionnelle de B sachant A** (en abrégé probabilité de B sachant A) et on note $\mathbf{P}_A(B)$ (ou parfois $\mathbf{P}(B|A)$) le réel

$$\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}.$$

Le plus dur ici est de bien comprendre que cette probabilité correspond à l'intuition qu'on s'en fait lorsque parle de probabilité de *B sachant A*.
Le cas qu'on comprend le mieux est celui d'équiprobabilité.

Exemple 27.14

Je suis convoqué pour faire passer un oral dans un lycée bien particulier, dont le site Web m'indique que tous les élèves sont en sport étude, soit en section escalade (26 élèves), soit en section bridge (47 élèves), et qu'il compte dans ses rangs 13 médaillés aux mondiaux junior (8 en escalade et 5 en bridge).
Avant le premier oral, je fais donc un petit calcul qui m'apprend que la probabilité que le prochain candidat soit un médaillé mondial est $\dfrac{13}{73}$.
J'ouvre la porte, le candidat entre. Ses biceps et son bronzage ne m'autorisent décemment pas à le penser en section bridge. Quelle est donc la probabilité qu'il s'agisse d'un médaillé mondial ? C'est-à-dire, la probabilité, *sachant que c'est un grimpeur*, qu'il s'agisse d'un médaillé mondial ?
La réponse est évidente, c'est $\dfrac{8}{26}$, le nombre de médaillés en escalade sur le nombre total de grimpeurs, mais si on réfléchit en termes de probabilités, $\dfrac{8}{26} = \dfrac{\frac{8}{77}}{\frac{26}{77}}$, où $\dfrac{8}{77}$ est la probabilité qu'un élève choisi au hasard soit **à la fois** grimpeur et médaillé, et $\dfrac{26}{77}$ est la probabilité qu'un élève choisi au hasard soit grimpeur. On trouve donc bien la formule $\dfrac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)}$.

⚠ La notation $\mathbf{P}(B|A)$ (qu'on essaiera autant que possible d'éviter, bien qu'elle figure au programme officiel) est trompeuse, et laisse penser qu'il existe un événement $B|A$, qui serait «B sachant A».
Un tel «événement conditionnel» **n'existe pas**, et lorsque vous aurez envie de manipuler de tels événements[13], c'est en fait bien souvent $A \cap B$ que vous aurez derrière la tête.

[13] Ce que vous vous abstiendrez bien entendu de faire !

Proposition 27.15 : *Si A est un événement tel que $\mathbf{P}(A) \neq 0$, alors l'application* $\mathbf{P}_A : \left| \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ B & \longmapsto & \mathbf{P}_A(B) \end{array} \right.$ *est une probabilité sur Ω.*

Une conséquence importante de ce fait est que toutes les règles énoncées pour le calcul des probabilités restent valables pour le calcul des probabilités conditionnelles !
Par exemple, $\mathbf{P}_A(\emptyset) = 0$, $\mathbf{P}_A\left(\overline{B}\right) = 1 - \mathbf{P}_A(B)$ et $\mathbf{P}_A(B \cup C) = \mathbf{P}_A(B) + \mathbf{P}_A(C) - \mathbf{P}_A(B \cap C)$.

Démonstration. Puisque la probabilité $\mathbf{P}$ est positive, croissante[14] et que $A \cap B \subset A$, on a bien pour tout $B \in \mathscr{P}(\Omega)$, $\mathbf{P}_A(B) = \dfrac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)} \in [0,1]$.
Par ailleurs, $\mathbf{P}_A(\Omega) = \dfrac{\mathbf{P}(A\cap \Omega)}{\mathbf{P}(A)} = \dfrac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1$.
Enfin, si B et C sont deux événements incompatibles, alors $A\cap B$ et $A \cap C$ le sont encore. Donc

$$\mathbf{P}_A(B\cup C) = \frac{\mathbf{P}(A\cap(B\cup C))}{\mathbf{P}(A)} = \frac{\mathbf{P}\left((A\cap B)\cup(A\cap C)\right)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A\cap B)+\mathbf{P}(A\cap C)}{\mathbf{P}(A)} = \mathbf{P}_A(B)+\mathbf{P}_A(C).$$

□

[14] Pour l'inclusion.

27.2.2 La formule des probabilités composées

Sur la définition de probabilité conditionnelle, il est assez clair que $\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}_A(B)$.
Ce résultat se généralise de la manière suivante :

Théorème 27.16 (Formule des probabilités composées) : *Soient $A_1, \dots, A_n$ des événements d'un espace probabilisé fini $(\Omega, \mathbf{P})$, tels que $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors*

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1)\mathbf{P}_{A_1}(A_2)\mathbf{P}_{A_1\cap A_2}(A_3)\cdots \mathbf{P}_{A_1\cap\dots\cap A_{n-1}}(A_n)$$
$$= \prod_{i=0}^{n-1} \mathbf{P}_{A_1\cap\dots\cap A_i}(A_{i+1}).$$

Remarque

La formule s'apprend et se comprend bien mieux avec les pointillés.

Démonstration. Notons que puisque pour tout $i \in [\![1, n-1]\!]$, $A_1 \cap \dots \cap A_{n-1} \subset A_1 \cap \dots \cap A_i$, et donc par croissance de la probabilité,

$$\mathbf{P}(A_1 \cap \dots \cap A_i) \geqslant \mathbf{P}(A_1 \cap \dots \cap A_{n-1}) > 0$$

de sorte que la probabilité conditionnelle $\mathbf{P}_{A_1\cap\dots\cap A_i}(A_{i+1})$ est bien définie[15].
Sinon, il s'agit de remarquer que le produit donné dans l'énoncé est télescopique :

$$\prod_{i=0}^{n-1} \mathbf{P}_{A_1\cap\dots\cap A_i}(A_{i+1})$$
$$= \cancel{\mathbf{P}(A_1)}\frac{\cancel{\mathbf{P}(A_1\cap A_2)}}{\cancel{\mathbf{P}(A_1)}}\frac{\cancel{\mathbf{P}(A_1\cap A_2\cap A_3)}}{\cancel{\mathbf{P}(A_1\cap A_2)}}\cdots\frac{\cancel{\mathbf{P}(A_1\cap\dots\cap A_{n-1})}}{\cancel{\mathbf{P}(A_1\cap\dots\cap A_{n-2})}}\frac{\mathbf{P}(A_1\cap\dots\cap A_n)}{\cancel{\mathbf{P}(A_1\cap\dots\cap A_{n-1})}}$$
$$= \mathbf{P}(A_1 \cap \dots \cap A_n).$$

□

[15] C'est d'ailleurs la seule raison d'être de cette hypothèse, qui ne pose en pratique aucun soucis.

Cette formule est absolument fondamentale lorsqu'on manipule des expériences successives telles que le résultat d'une expérience influe sur la suivante. L'exemple typique étant les tirages successifs **sans remise**.

Exemple 27.17

On effectue des tirages sans remise dans une urne qui contient 2 boules rouges et $n-2$ boules vertes.
Pour $k \in [\![1, n-1]\!]$, on note A_k l'événement «la première boule rouge sort au $k^{\text{ème}}$ tirage».
Notons également R_i l'événement «la boule obtenue au $i^{\text{ème}}$ tirage est rouge».
Nous ne disposons pas directement des probabilités des R_i, mais plutôt de leurs probabilités conditionnelles connaissant les résultats des tirages précédents !
On a $A_k = \left(\bigcap_{i=1}^{k-1} \overline{R_i}\right) \cap R_k$.
Et donc

$$\begin{aligned}\mathbf{P}(A_k) &= \mathbf{P}\left(\left(\bigcap_{i=1}^{k-1} \overline{R_i}\right) \cap R_k\right)\\ &= \mathbf{P}(\overline{R_1})\mathbf{P}_{\overline{R_1}}(\overline{R_2})\mathbf{P}_{\overline{R_1}\cap\overline{R_2}}(\overline{R_3})\cdots\mathbf{P}_{\overline{R_1}\cap\dots\cap\overline{R_{k-2}}}(\overline{R_{k-1}})\mathbf{P}_{\overline{R_1}\cap\dots\cap\overline{R_{k-1}}}(R_k)\\ &= \frac{n-2}{n}\frac{n-3}{n-1}\frac{n-4}{n-2}\cdots\frac{n-2-(k-2)}{n-(k-2)}\frac{2}{n-(k-1)}\\ &= \frac{2(n-k)}{n(n-1)}.\end{aligned}$$

27.2.3 La formule des probabilités totales

La formule des probabilités totales sert tout le temps, et vous l'avez déjà beaucoup utilisée, en vous gardant bien de le dire : elle sert précisément à faire des distinctions de cas afin de se ramener à des cas où les calculs de probabilités sont plus simples, puis de «remettre ensemble» toutes les probabilités obtenues.

Théorème 27.18 (Formule des probabilités totales) : *Soit $\{A_1, \dots, A_n\}$ un système complet d'événements. Alors pour tout $B \in \mathscr{P}(\Omega)$,*

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

Si de plus, tous les A_i sont de probabilités non nulles, alors

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

Méthode

Il faut connaître les deux formes données ici, et être capable de décider laquelle on souhaite utiliser dans une situation concrète. Ce qui dépend essentiellement de ce que vous pouvez calculer le plus facilement : des probabilités d'intersection ou des probabilités conditionnelles ?

Démonstration. Puisque $\Omega = \bigcup_{i=1}^{n} A_i$, et que ces événements sont deux à deux incompatibles, alors $B = B \cap \Omega = \bigcup_{i=1}^{n} (B \cap A_i)$, et les $B \cap A_i$ sont deux à deux incompatibles.

Donc il vient $\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(B \cap A_i)$.

La seconde forme, dans le cas où les $\mathbf{P}(A_i)$ sont tous non nuls s'en déduit immédiatement en notant que $\mathbf{P}(B \cap A_i) = \mathbf{P}(A_i)\mathbf{P}_{A_i}(B)$. □

Si certains des $\mathbf{P}(A_i)$ sont nuls, ce qui normalement ne devrait pas trop arriver sur un univers fini, on a

$$\mathbf{P}(B) = \sum_{i \mid \mathbf{P}(A_i) \neq 0} \mathbf{P}(A_i)\mathbf{P}_{A_i}(B).$$

Exemples 27.19

▶ On lance 3 fois une pièce qui tombe sur PILE avec probabilité p et sur FACE avec probabilité $q = 1 - p$.
On note P_i (resp. F_i) l'événement «le $i^{\text{ème}}$ lancer donne PILE (resp. FACE)». On note A l'événement «deux lancers consécutifs n'ont jamais donné le même résultat».
Utilisons alors la formule des probabilités totales à l'aide du système complet d'événements $\{P_2, F_2\}$.

$$\mathbf{P}(A) = \mathbf{P}(A \cap P_2) + \mathbf{P}(A \cap F_2) = \mathbf{P}(F_1 \cap P_2 \cap F_3) + \mathbf{P}(P_1 \cap F_2 \cap P_3) = q^2 p + p^2 q.$$

Remarque

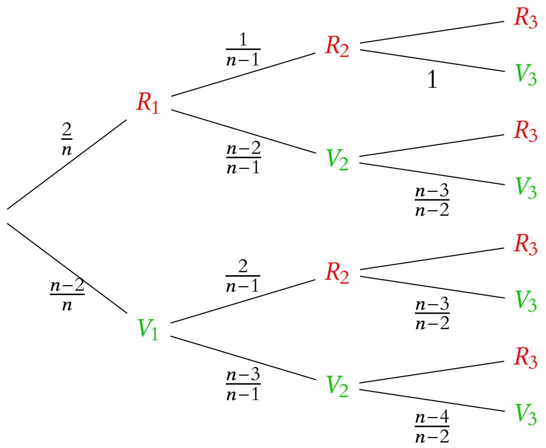
La véritable justification des produits, qui vient de l'indépendance des lancers, viendra un peu plus tard.

▶ Une urne contient $n \geqslant 3$ boules : 2 boules rouges et $n-2$ boules vertes. On tire successivement et sans remise 3 boules. Quelle est la probabilité que la troisième boule soit verte ?
Avec des notations évidentes, $\{R_1 \cap R_2, R_1 \cap V_2, V_1 \cap R_2, V_1 \cap V_2\}$ est un système complet d'événements.
On a alors $\mathbf{P}(R_1 \cap R_2) = \mathbf{P}(R_1)\mathbf{P}_{R_1}(R_2) = \dfrac{2}{n}\dfrac{1}{n-1}$.
Et de même,

$$\mathbf{P}(R_1 \cap V_2) = \frac{2}{n}\frac{n-2}{n-1},\ \mathbf{P}(V_1 \cap R_2) = \frac{n-2}{n}\frac{2}{n-1},\ \mathbf{P}(V_1 \cap V_2) = \frac{n-2}{n}\frac{n-3}{n-1}.$$

Donc par la formule des probabilités totales, avec le système complet d'événements susmentionné

$$\begin{aligned}
\mathbf{P}(V_3) &= \mathbf{P}(R_1 \cap R_2)\mathbf{P}_{R_1 \cap R_2}(V_3) + \mathbf{P}(R_1 \cap V_2)\mathbf{P}_{R_1 \cap V_2}(V_3) + \mathbf{P}(V_1 \cap R_2)\mathbf{P}_{V_1 \cap R_2}(V_3) + \mathbf{P}(V_1 \cap V_2)\mathbf{P}_{V_1 \cap V_2}(V_3) \\
&= \frac{2}{n(n-1)} \times 1 + \frac{2(n-2)}{n(n-1)}\frac{n-3}{n-2} + \frac{2(n-2)}{n(n-1)}\frac{n-3}{n-2} + \frac{(n-2)(n-3)}{n(n-1)}\frac{n-4}{n-2} \\
&= \frac{1}{n(n-1)} \times (2 + 4(n-3) + (n-3)(n-4)) \\
&= \frac{n^2 - 3n + 2}{n(n-1)} = \frac{n-2}{n}.
\end{aligned}$$



La formule des probabilités totales et la formule des probabilités composées sont les deux fondements des arbres de probabilités que vous avez pu utiliser au lycée : les probabilités indiquées sur les branches sont des probabilités conditionnelles, la formule des probabilités composées justifie qu'on fasse le produit des probabilités situées sur des branches qui se suivent, et la formule des probabilités totales justifie qu'on finisse par sommer toutes les probabilités aboutissant à la réalisation d'un événement donné (ici V_3).

► On dispose de $n+1$ urnes $U_0, U_1, \dots, U_n$, qui contiennent chacune n boules : l'urne U_k contient k boules bleues et $n-k$ boules rouges.
On choisit une urne au hasard (et de manière équiprobable), et on tire simultanément deux boules dans cette urne. Quelle est la probabilité que les deux boules soient bleues ?

Notons[16] $U_0, \dots, U_n$ les événements «le tirage a lieu dans l'urne U_i», de sorte que $\{U_0, \dots, U_n\}$ est un système complet d'événements.
Notons également A l'événement «les deux boules sont bleues».
Il est évident que la probabilité que A se réalise dépend de l'urne dans laquelle le tirage a lieu.
Par la formule des probabilités totales appliquée avec le système complet d'événements $\{U_0, \dots, U_n\}$, il vient

$$\mathbf{P}(A) = \sum_{k=0}^{n} \mathbf{P}(U_k)\mathbf{P}_{U_k}(A) = \sum_{k=0}^{n} \frac{1}{n+1}\mathbf{P}_{U_k}(A).$$

Or, il est facile de constater que $\mathbf{P}_{U_0}(A) = \mathbf{P}_{U_1}(A) = 0$.
Et pour $k \geqslant 2$, alors l'urne U_k contient n boules, donc il y a $\binom{n}{2}$ manières d'y tirer 2 boules, et puisqu'il y a k boules bleues, il y a $\binom{k}{2}$ manières d'y tirer 2 boules bleues.

Autrement dit $\mathbf{P}_{U_k}(A) = \dfrac{\binom{k}{2}}{\binom{n}{2}} = \dfrac{k(k-1)}{n(n-1)}$.

Et donc

$$\begin{aligned}\mathbf{P}(A) &= \frac{1}{n+1}\sum_{k=2}^{n}\frac{k(k-1)}{n(n-1)} = \frac{1}{(n+1)n(n-1)}\sum_{k=1}^{n}k(k-1)\\ &= \frac{1}{(n+1)n(n-1)}\left(\sum_{k=1}^{n}k^2 - \sum_{k=1}^{n}k\right)\\ &= \frac{1}{(n+1)n(n-1)}\left(\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}\right)\\ &= \frac{1}{n-1}\,\frac{2n-2}{6} = \frac{1}{3}.\end{aligned}$$

Le terme $k=1$ est nul.

Grow up !
Si un arbre est un excellent support au brouillon, il ne peut plus en aucun cas être considéré comme une preuve, et vous devez justifier tous vos calculs !

[16] Abusivement.

⚠ Attention !
Pour autant $\mathbf{P}(A)$ n'a qu'une et qu'une seule valeur.
Je ne veux surtout pas voir de «si le tirage a lieu dans l'urne i, alors $\mathbf{P}(A) = \dots$.
Ce qu'on donnerait ainsi, ce sont des probabilités conditionnelles $\mathbf{P}_{U_i}(A)$.
La formule des probabilités totales est justement le moyen de rassembler toutes ces probabilités pour déterminer **la** valeur de $\mathbf{P}(A)$.

27.2.4 Les formules de Bayes

Proposition 27.20 (Formule de Bayes) : *Soient A et B deux événements de probabilités non nulles. Alors*
$$\mathbf{P}_B(A) = \frac{\mathbf{P}_A(B)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

Démonstration.
$$\mathbf{P}_B(A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}_A(B)\mathbf{P}(A)}{\mathbf{P}(B)}.$$
□

Exemple 27.21

Les situations où il y a besoin de la formule de Bayes sont assez facilement reconnaissables : ce sont celles où on cherche une probabilité conditionnelle $\mathbf{P}_B(A)$, mais où on connaît la probabilité conditionnelle «dans l'autre sens» $\mathbf{P}_A(B)$.
Reprenons par exemple les urnes $U_0, \dots, U_n$ de l'exemple 27.19.
On a tiré deux boules, et les deux sont bleues. Quelle est la probabilité que le tirage ait eu lieu dans l'urne U_k ?
Avec les mêmes notations que précédemment, il s'agit donc de calculer $\mathbf{P}_A(U_k)$. Par la formule de Bayes,
$$\mathbf{P}_A(U_k) = \frac{\mathbf{P}(U_k)\mathbf{P}_{U_k}(A)}{\mathbf{P}(A)} = \frac{\frac{1}{n+1}\frac{k(k-1)}{n(n-1)}}{\frac{1}{3}} = \frac{3k(k-1)}{(n+1)n(n-1)}.$$
Un bon moyen de vérifier le résultat est souvent de regarder ce qui se passe pour les valeurs extrêmes. Ici, on sait que si le tirage a donné deux boules bleues, il ne peut avoir eu lieu ni dans l'urne U_0, ni dans l'urne U_1.
Or notre formule donne une probabilité nulle pour $k \in \{0, 1\}$, ce qui est cohérent.
De plus, le fait que la probabilité soit une fonction croissante de k est aussi conforme à l'intuition.

Une autre forme parfois utile de la formule de Bayes est la suivante :

Corollaire 27.22 – *Soit $\{A_1, \dots, A_n\}$ un système complet d'événements de probabilités non nulles, et soit B un événement de probabilité non nulle. Alors*
$$\forall j \in [\![1, n]\!],\ \mathbf{P}_B(A_j) = \frac{\mathbf{P}_{A_j}(B)\mathbf{P}(A_j)}{\displaystyle\sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{P}_{A_i}(B)}.$$

> **Remarque**
> Je la mentionne parce qu'elle est au programme, mais comme il s'agit simplement de coupler la formule de Bayes ci-dessus à la formule des probabilités totales, normalement vous saurez toujours la retrouver dans les cas où vous en aurez besoin, sans qu'il soit nécessaire de connaître un énoncé général.

Démonstration. Il suffit juste de partir de la formule de Bayes, avec $A = A_j$, et d'appliquer la formule des probabilités totales au dénominateur afin de calculer $\mathbf{P}(B)$. □

27.3 INDÉPENDANCE D'ÉVÉNEMENTS

Définition 27.23 – Deux événements A et B d'un même espace probabilisé fini $(\Omega, \mathbf{P})$ sont dits **indépendants** si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

Notons bien que ceci fait appel à la probabilité $\mathbf{P}$, et ne dépend donc pas que de l'univers Ω, mais bien de la probabilité $\mathbf{P}$ que l'on met dessus.
La plupart du temps, l'indépendance est supposée par l'énoncé, ou alors est implicite, notamment lors de la répétition d'expériences telles que l'issue d'une expérience n'influe pas sur les suivantes. Typiquement lors de plusieurs lancers d'une pièce/d'un dé/whatever ou

de tirages successifs **avec remise.**

> Rédaction
> Dans ces exemples, un énoncé ne mentionnera jamais l'indépendance des répétitions, ce sera à vous de la comprendre et de la mentionner si vous l'utilisez.

⚠ En dehors des exemples ci-dessus, ne vous fiez pas trop à l'intuition pour l'indépendance, c'est une notion parfois imprévisible qui peut contredire l'intuition.

Exemple 27.24

On lance deux fois une pièce de monnaie qui tombe sur FACE avec probabilité $p \in [0, 1]$.
Notons P_i (resp. F_i) l'événement «le $i^{\text{ème}}$ lancer donne PILE», et A l'événement «les deux lancers donnent le même résultat».
Alors $\mathbf{P}(F_1) = p$ et

$$\mathbf{P}(A) = \mathbf{P}\left((F_1 \cap F_2) \cup (P_1 \cap P_2)\right) = \mathbf{P}(F_1 \cap F_2) + \mathbf{P}(P_1 \cap P_2) = \mathbf{P}(F_1)\mathbf{P}(F_2) + \mathbf{P}(P_1)\mathbf{P}(P_2) = p^2 + (1-p)^2.$$

Enfin, on a $\mathbf{P}(A \cap F_1) = \mathbf{P}(F_1 \cap F_2) = p^2$.
On a alors

$$\mathbf{P}(A)\mathbf{P}(F_1) = \mathbf{P}(A \cap F_1) \Leftrightarrow p^3 + p(1-p)^2 = p^2 \Leftrightarrow p\left(p^2 + 1 - 2p + p^2 - p\right) = 0 \Leftrightarrow p(2p^2 - 3p + 1) = 0 \Leftrightarrow p \in \left\{0, \frac{1}{2}, 1\right\}.$$

Donc pour $p \notin \left\{0, \frac{1}{2}, 1\right\}$, les événements A et F_1 ne sont pas indépendants, alors qu'ils le sont dans le cas contraire.
Pourtant, dans les deux cas, l'univers considéré est le même : c'est $\{P, F\}^2$.

Notons que deux événements A et B incompatibles et de probabilités non nulles ne peuvent pas être indépendants, puisqu'alors $\mathbf{P}(A \cap B) = 0 \neq \mathbf{P}(A)\mathbf{P}(B)$.
C'est souvent là un bon moyen de prouver la non-indépendance de deux événements.

Proposition 27.25 : *Deux événements A et B avec $\mathbf{P}(A) \neq 0$ sont indépendants si et seulement si $\mathbf{P}_A(B) = \mathbf{P}(B)$.*

Démonstration. C'est évident puisque $\mathbf{P}_A(B) = \dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}$ et donc

$$\mathbf{P}_A(B) = \mathbf{P}(B) \Leftrightarrow \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A \cap B).$$

□

L'intuition là-derrière est que si A et B sont indépendants, alors savoir que A est réalisé ne change rien à la probabilité que B le soit.

Le cas où $\mathbf{P}(A) = 0$ est peu intéressant, car un tel événement[17] est indépendant de tout autre événement B, puisque $A \cap B \subset A$ et donc $\mathbf{P}(A \cap B) \leqslant \mathbf{P}(A) = 0$, de sorte que $\mathbf{P}(A \cap B) = 0 = \mathbf{P}(A)\mathbf{P}(B)$.

[17] De probabilité nulle signifiant qu'il ne se produit jamais ou presque...

Proposition 27.26 : *Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants.*

> Corollaire
> Et donc $\overline{A}$ et $\overline{B}$ sont également indépendants, de même que $\overline{A}$ et B.

Démonstration. Puisque $\{B, \overline{B}\}$ est un système complet d'événements, par la formule des probabilités totales, $\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B})$.
Et donc

$$\mathbf{P}(A \cap \overline{B}) = \mathbf{P}(A) - \mathbf{P}(A \cap B) = \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A)\left(1 - \mathbf{P}(B)\right) = \mathbf{P}(A)\mathbf{P}(\overline{B})$$

de sorte que A et $\overline{B}$ sont bien indépendants. □

Définition 27.27 – Des événements $A_1, \dots, A_n$ d'un espace probabilisé $(\Omega, \mathbf{P})$ sont dits **mutuellement indépendants** si $\forall I \subset [\![1, n]\!]$, $\mathbf{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbf{P}(A_i)$.

⚠ Pour $n \geqslant 3$, l'indépendance mutuelle de n événements implique bien qu'ils sont deux à deux indépendants (il suffit de prendre I de cardinal 2 dans la définition), mais la réciproque est fausse.
Par exemple, lançons deux dés, et notons A «le numéro du premier dé est pair», B «le numéro du second dé est pair» et C «la somme des deux dés est paire».
Alors on a $\mathbf{P}(A) = \mathbf{P}(B) = \mathbf{P}(C) = \dfrac{1}{2}$.
Par ailleurs, il est clair que A et B sont indépendants, et

$$\mathbf{P}(A \cap C) = \mathbf{P}(A \cap B) = \frac{1}{4}.$$

Donc A, B, C sont deux à deux indépendants.
En revanche, $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A \cap B) = \dfrac{1}{4}$, alors que $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \dfrac{1}{8}$.
Donc A, B et C ne sont pas mutuellement indépendants.

Détails
Pour obtenir $\mathbf{P}(C)$, le plus simple est probablement de procéder par dénombrement. Il y a 36 résultats possibles (et équiprobables) à l'expérience, dont

$$1 + 3 + 5 + 5 + 3 + 1 = 18$$

rendent la somme des deux dés paire.

Détails
Si A et B sont réalisés, alors C l'est.
Autrement dit, $A \cap B \subset C$.

27.4 VARIABLES ALÉATOIRES

27.4.1 Définition

Difficile
Il n'est pas toujours évident de comprendre en quoi une fonction (quelque chose qui est on ne peut plus déterministe) nous aide à étudier les phénomènes aléatoires. C'est effectivement surprenant au premier abord, mais un passage obligé...

Définition 27.28 – Soit Ω un ensemble fini, et soit E un ensemble. On appelle **variable aléatoire sur Ω à valeurs dans E** toute application $X : \Omega \to E$.
Dans le cas où $E = \mathbf{R}$, on parle de **variable aléatoire réelle.**

Pour l'instant, la notion de variable aléatoire n'a évidemment pas l'air d'un grand intérêt, puisque toute application sur Ω est une variable aléatoire, mais un peu de patience. Notons qu'on n'a même pas besoin de mettre une probabilité sur $\mathbf{P}$ pour parler d'une variable aléatoire. Mais là aussi patience, les variables aléatoires prendront tout leur intérêt une fois qu'on les couplera à des probabilités.

Profitons-en pour faire la différence entre les événements et les variables aléatoires. Un événement est un ensemble d'événements élémentaires, c'est-à-dire d'issues de l'expérience. Par exemple[18] lorsqu'on lance deux dés, et qu'on modélise ceci par $\Omega = [\![1, 6]\!]^2$, l'événement A : «la somme des deux dés vaut 3» est $\{(1, 2), (2, 1)\}$.
Une variable aléatoire associe un nombre[19] à **chaque** issue de l'expérience.
Par exemple, la variable aléatoire «somme des deux dés» est définie quel que soit l'issue de l'expérience. Cette variable aléatoire n'est rien d'autre que $X : \left|\begin{array}{ccc} [\![1,6]\!]^2 & \longrightarrow & \mathbf{R} \\ (i,j) & \longmapsto & i+j \end{array}\right.$.
Il y a bien un lien avec des événements, que nous allons expliciter dans un instant, par exemple l'événement A ci-dessus n'est rien d'autre que l'événement «la variable X vaut 3». Mais il est important de bien comprendre la différence de nature entre une variable aléatoire et un événement. Notamment pour savoir ce qu'on peut écrire et ce qu'on ne peut pas écrire. Par exemple, un événement est réalisé ou ne l'est pas, ce qui n'a pas de sens pour une variable aléatoire. Et on ne parle donc pas du contraire d'une variable aléatoire.

Sur le même espace probabilisé existent d'autres variables aléatoires, par exemple «le plus grand numéro des deux dés», qui est $Y : \left|\begin{array}{ccc} [\![1,6]\!]^2 & \longrightarrow & \mathbf{R} \\ (i,j) & \longmapsto & \max(i,j) \end{array}\right.$, ou encore le numéro du second dé qui est $Z : \left|\begin{array}{ccc} [\![1,6]\!]^2 & \longrightarrow & \mathbf{R} \\ (i,j) & \longmapsto & j \end{array}\right.$.

Pour reprendre un exemple de l'introduction, lorsqu'on lance n fois une pièce (que l'on modélise par $\{P, F\}^n$), la variable aléatoire «le numéro du premier PILE» n'est pas bien définie car on ne lui attribue pas de valeur dans le cas où on n'obtient que des FACES.

[18] Et c'est de loin mon *toy model* préféré, à utiliser sans modération lorsque vous avez un doute sur la signification des notions de base en probas.

[19] Au moins dans le cas d'une variable aléatoire réelle.

En revanche, $X : \left| \begin{array}{ccl} \Omega & \longrightarrow & \mathbf{N} \cup \{+\infty\} \\ (a_1, \dots, a_n) & \longmapsto & \begin{cases} +\infty & \text{si } (a_1, \dots, a_n) = (F, F, \dots, F) \\ \min\{i \in [\![1, n]\!] \mid a_i = P\} & \text{sinon} \end{cases} \end{array} \right.$

est bien une variable aléatoire.

Donnons un exemple de variable aléatoire qui n'est pas une variable aléatoire réelle, mais dans la suite nous ne parlerons quasiment que de variables à valeurs réelles.

Si on tire sans remise deux boules dans une urne qui contient 3 boules rouges, 2 noires et une blanche, on peut considérer la variable aléatoire à valeurs dans $\{R, N, B\}$ qui donne la couleur de la seconde boule.

À l'aide d'une variable aléatoire, il est possible de former de nombreux événements, et c'est là le premier intérêt des variables aléatoires : nous aider à nommer facilement des événements.

Définition 27.29 – Soit $X : \Omega \to E$ une variable aléatoire. Pour $A \subset E$, on note $[X \in A]$ l'événement[20] défini par

$$[X \in A] := X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}.$$

Lorsque $A = \{a\}$, on note $[X = a]$ au lieu de $[X \in \{a\}]$.
Et lorsque X est une variable aléatoire réelle, on note $[X \leqslant a] = X^{-1}(]-\infty, a])$, $[X < a] = X^{-1}(]-\infty, a[)$, et de même, on utilise les notations $[X \geqslant a]$, $[X > a]$, $[a \leqslant X \leqslant b] = X^{-1}([a, b])$, etc.

[20] C'est une partie de Ω.

Exemple 27.30

Toujours en lançant deux dés, notons X le résultat du premier dé, Y le résultat du second, et $Z = X + Y$ la somme des deux.
Alors l'événement «le premier dé vaut 1» est $[X \leqslant 1]$, l'événement «le second dé est pair» est $[Y = 2] \cup [Y = 4] \cup [Y = 6]$.
Et l'événement $[Z > 10]$ est $\{(5, 6), (6, 5), (6, 6)\}$.

⚠ Entendons-nous bien : un événement peut avoir une probabilité, pas une variable aléatoire !
Donc si X est une variable aléatoire, $\mathbf{P}([X = 2])$, $\mathbf{P}([X \geqslant 3])$ et $\mathbf{P}([X \geqslant 3] \cup [X < -2])$ ont un sens. En revanche, $\mathbf{P}(X)$ **ne veut rien dire.**
Pas plus que $\mathbf{P}([X \geqslant 3]) \cup \mathbf{P}([X < -2])$ ou $\mathbf{P}([X \geqslant 3] + [X < -2])$, mais c'est une autre histoire[21].

Rédaction
En général on oublie les crochets quand on écrit des probas : $\mathbf{P}(X \leqslant a)$ et pas $\mathbf{P}([X \leqslant a])$.

[21] On peut considérer l'union d'événements, ou la somme de probas, pas le contraire !

Définition 27.31 – Si $X : \Omega \to E$ est une variable aléatoire, alors son image $X(\Omega)$ (qu'il faut interpréter comme l'ensemble des valeurs prises par X) est appelé **support de X.**
Puisque Ω est fini, il s'agit toujours d'un ensemble fini.

Proposition 27.32 : *Soit X une variable aléatoire sur Ω. Alors $\{[X = x], x \in X(\Omega)\}$ est un système complet d'événements, appelé système complet d'événements associé à X.*

Démonstration. Il suffit de se rappeler que $[X = x] = \{\omega \in \Omega \mid X(\omega) = x\}$.
Il est clair que les $[X = x]$ sont deux à deux incompatibles[22].
Et puisque $X(\Omega)$ est l'image de X, on a bien[23]

$$\bigcup_{x \in X(\Omega)} [X = x] = \Omega.$$

[22] Disjoints.
[23] Tout élément de Ω a son image dans $X(\Omega)$.

□

L'idée est tout simplement qu'on distingue plusieurs cas suivant la valeur que prend X, et comme elle prend toujours une valeur, on a bien ainsi distingué tous les cas.
Un cas très particulier de variable aléatoire est celle de l'indicatrice d'un événement A, qui rappelons-le, est définie par $\mathbb{1}_A : \left| \begin{array}{ccc} \Omega & \longrightarrow & \{0,1\} \\ \omega & \longmapsto & \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \end{array} \right.$.
Autrement dit, c'est la variable aléatoire qui vaut 1 si A est réalisé et 0 sinon.
Et on a donc $A = [\mathbb{1}_A = 1]$.

Exemple 27.33

Toujours en lançant deux dés, on note A l'événement «les deux dés ont même parité».
Alors $\mathbb{1}_A$ est la variable aléatoire qui vaut 1 si et seulement si les deux dés ont même parité.
Donc $[\mathbb{1}_A = 1] = \{(1,3),(3,1),(1,5),(2,6),\dots\}$.

27.4.2 Loi d'une variable aléatoire

Venons-en enfin aux probabilités : lorsque Ω est muni d'une probabilité $\mathbf{P}$, on peut parler de probabilité d'un événement, et en particulier de tous les événements qui sont associés à une variable aléatoire.
Par exemple, lorsqu'on joue au tiercé, en notant X le numéro du cheval gagnant, pour parier intelligemment il faut connaître les $\mathbf{P}([X=1]), \mathbf{P}([X=2])$, etc. La donnée de toutes ces probabilités est ce qu'on appelle la loi de X.

> **Cote/probas**
> La probabilité qu'un cheval gagne est grosso modo (c'est-à- dire modulo la commission du bookmaker) l'inverse de sa cote.

Définition 27.34 – Soit X une variable aléatoire sur $(\Omega, \mathbf{P})$. La **loi de** X est alors l'application $\mathbf{P}_X : \left| \begin{array}{ccc} \mathscr{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbf{P}(X \in A) = \mathbf{P}(X^{-1}(A)) \end{array} \right.$.

Proposition 27.35 : *La loi* $\mathbf{P}_X$ *de* X *est une probabilité sur* $X(\Omega)$.

Démonstration. Il est clair que $\mathbf{P}_X$ est à valeurs dans $[0,1]$, et

$$\mathbf{P}_X(X(\Omega)) = \mathbf{P}\left(X^{-1}(X(\Omega))\right) = \mathbf{P}(\Omega) = 1.$$

Si A et B sont deux parties disjointes de $X(\Omega)$, alors

$$\mathbf{P}_X(A \cup B) = \mathbf{P}_X(A \cup B) = \mathbf{P}\left(X^{-1}(A \cup B)\right) = \mathbf{P}\left(X^{-1}(A) \cup X^{-1}(B)\right)$$

et ces deux événements étant incompatibles,

$$\mathbf{P}_X(A \cup B) = \mathbf{P}(X^{-1}(A)) + \mathbf{P}(X^{-1}(B)) = \mathbf{P}_X(A) + \mathbf{P}_X(B).$$

□

La loi d'une variable aléatoire dépend de la probabilité qu'on utilise sur Ω.
Par exemple, si on lance deux fois un dé, et qu'on note X le résultat du premier lancer, alors que le dé soit équilibré ou non, on peut modéliser l'expérience par $\Omega = [\![1,6]\!]^2$. Et alors X est toujours l'application $(i,j) \mapsto i$.
Par contre, si le dé est équilibré, on va utiliser la probabilité uniforme sur $[\![1,6]\!]$ et on vérifie alors que $\mathbf{P}_X$ est la probabilité uniforme sur $[\![1,6]\!]$.
En revanche, si le dé est pipé, on va utiliser un autre $\mathbf{P}$, ce qui donnera une autre loi pour X.

Proposition 27.36 : *Soit X une variable aléatoire sur Ω. Alors la loi de X est entièrement déterminée par les $\mathbf{P}_X(\{x\}) = \mathbf{P}(X = x)$, $x \in X(\Omega)$.*
Plus précisément, pour $A \in \mathscr{P}(X(\Omega))$, on a

$$\mathbf{P}_X(A) = \sum_{x \in A} \mathbf{P}(X = x).$$

Démonstration. Il suffit de noter que pour $A \in \mathscr{P}(X(\Omega))$, $A = \bigcup_{x \in A} \{x\}$ et donc

$$\mathbf{P}_X(A) = \mathbf{P}\left(X^{-1}(A)\right) = \mathbf{P}\left(X^{-1}\left(\bigcup_{x \in A}\{x\}\right)\right) = \mathbf{P}\left(\bigcup_{x \in A} X^{-1}(\{x\})\right) = \sum_{x \in A} \mathbf{P}_X(\{x\}) = \sum_{x \in A} \mathbf{P}(X = x).$$

□

Ainsi, lorsqu'on déterminera la loi d'une variable aléatoire X, on n'explicitera pas la probabilité $\mathbf{P}_X$ (et heureusement car il s'agit d'une application définie sur $\mathscr{P}(X(\Omega))$), mais on se contentera de déterminer :

- son support $X(\Omega)$
- les $\mathbf{P}(X = x)$, pour $x \in X(\Omega)$.

Autrement dit, on donnera les valeurs prises par la variable, et les probabilités qu'elle prenne chacune de ces valeurs.

Exemple 27.37

Une urne contient N boules numérotées de 1 à N.
On tire simultanément n ($n \leqslant N$) boules dans cette urne, et on note X le plus grand numéro ainsi obtenu.
Alors $X(\Omega) = [\![n, N]\!]$, et pour déterminer la loi de X, il nous faut donner en plus les $\mathbf{P}(X = k)$, $k \in [\![n, N]\!]$.
Il y a en tout $\binom{N}{n}$ tirages possibles, tous équiprobables. Parmi ceux-ci, ceux qui donnent une plus grande boule portant le numéro k sont au nombre de $\begin{pmatrix} k-1 \\ n-1 \end{pmatrix}$. En effet, un tel tirage est formé de la boule numéro k, et de $n - 1$ boules de numéro strictement inférieur à k.
Donc $\mathbf{P}(X = k) = \dfrac{\binom{k-1}{n-1}}{\binom{N}{n}}$.

⚠ Deux variables définies sur un même espace et de même loi ne sont pas égales pour autant.
Prenons encore et toujours le lancer de deux dés équilibrés, et notons X le résultat du premier dé, et Y le résultat du second.
Alors X et Y ont même loi (que nous nommerons forme), mais ne sont pas égales, par exemple car $X(1, 2) = 1$ alors que $Y(1, 2) = 2$.
De même, si on lance n fois une pièce équilibrée, et que l'on note X (resp. Y) le nombre de PILES (resp. de FACES), alors X et Y ont même loi[24] sans pour autant être égales.

[24] Contentons-nous de l'affirmer, mais cela doit vous sembler intuitif.

Puisque $\{[X = x], x \in X(\Omega)\}$ est un système complet d'événements, on a évidemment $\sum_{x \in X(\Omega)} \mathbf{P}(X = x) = 1$.
Il existe une réciproque à ce résultat.

Déjà vu ?
C'est le sacro-saint «la somme des probas vaut 1».

Proposition 27.38 : *Soit $n \geqslant 1$, soient $x_1, \dots, x_n$ des éléments deux à deux distincts d'un ensemble E, et soient $p_1, \dots, p_n$ des réels positifs, tels que $\sum_{i=1}^{n} p_i = 1$.*
Alors il existe un ensemble fini Ω, une probabilité $\mathbf{P}$ sur Ω et une variable aléatoire X définie sur Ω et à valeurs dans E telle que

$$\forall i \in [\![1, n]\!],\ \mathbf{P}(X = x_i) = p_i.$$

Remarque
Les p_i étant positifs et de somme 1, ils sont automatiquement dans $[0, 1]$.

Démonstration. Prouvons un résultat plus fort, à savoir que sur tout ensemble fini Ω de cardinal supérieur ou égal à n, il existe une probabilité $\mathbf{P}$ et une variable aléatoire X définie sur Ω telle que $\forall i \in [\![1, n]\!]$, $\mathbf{P}(X = x_i) = p_i$.
Soit donc Ω un ensemble de cardinal $\geqslant n$ et soient $\omega_1, \dots, \omega_n$ n éléments distincts de Ω.
Alors par la proposition 27.12, il existe[25] une (unique) probabilité $\mathbf{P}$ sur Ω telle que

$$\mathbf{P}(\{\omega\}) = \begin{cases} p_i & \text{si } \omega = \omega_i \\ 0 & \text{si } \omega \notin \{\omega_1, \dots, \omega_n\} \end{cases}$$

Et alors en posant $X : \left| \begin{array}{ccc} \Omega & \longrightarrow & E \\ \omega & \longmapsto & \begin{cases} x_i & \text{si } \omega = \omega_i \\ 0 & \text{sinon} \end{cases} \end{array} \right.$, alors X est une application définie sur Ω, donc une variable aléatoire sur Ω.
Et on a alors clairement $X(\Omega) = \{x_1, \dots, x_n\}$, et pour tout $i \in [\![1, n]\!]$,

$$\mathbf{P}(X = x_i) = \mathbf{P}(X^{-1}(\{x_i\})) = \mathbf{P}(\{\omega_i\}) = p_i.$$

□

[25] Car $p_1 + \dots + p_n = 1$.

Cette proposition est en réalité très importante, sans qu'il soit vraiment nécessaire d'en maîtriser l'énoncé[26]. En effet, elle nous dispense, lorsqu'on manipule des variables aléatoires, de s'interroger sur l'espace probabilisé $(\Omega, \mathbf{P})$ que l'on utilise.
Par exemple, si on souhaite modéliser le lancer d'un dé pipé qui tombe sur 1 avec probabilité $1/4$, sur 2 avec probabilité $1/8$, sur 3 avec probabilité $1/8$ et sur 4, 5 ou 6 avec probabilité $1/6$ chacun, nul besoin de construire explicitement un univers Ω et une probabilité $\mathbf{P}$ sur laquelle une variable aléatoire dont la loi est celle que l'on veut, il suffit de vérifier que $\frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$, ce qui est bien le cas.

[26] Dont vous n'aurez jamais besoin.

C'est la raison pour laquelle beaucoup d'exercices commencent par «soit X une variable aléatoire suivant la loi *blabla*[27]». On sait qu'alors il existe un espace probabilisé sur lequel ceci a bien un sens, et cela nous suffit, nous n'avons pas besoin d'en savoir plus au sujet de cet espace.
De manière générale, la connaissance de l'univers Ω ne peut avoir d'intérêt que dans les cas où on peut faire du dénombrement, c'est-à-dire quand il y a équiprobabilité.

[27] Par exemple une loi uniforme ou une loi binomiale, voir ci-dessous.

Enfin, il faudra bien comprendre que les énoncés du type : «montrer que les réels $(p_i)_{1 \leqslant i \leqslant n}$ définissent bien une loi de probabilité sur $\{x_1, \dots, x_n\}$» signifient juste «prouver que les p_i sont des réels positifs de somme 1».

Exemple 27.39

À quelle condition sur $\lambda \in \mathbf{R}$ existe-t-il une variable aléatoire X à valeurs dans $[\![1, n]\!]$ telle que $\forall i \in [\![1, n]\!]$, $\mathbf{P}(X = i) = \lambda i$?
Il faut bien évidemment que λ soit positif pour que les λi le soient, et il faut de plus que

$$\sum_{i=1}^{n} \lambda i = 1 \Leftrightarrow \lambda = \frac{2}{n(n+1)}.$$

Puisque ce nombre est bien positif, on en déduit que $\dfrac{2}{n(n+1)}$ est la seule solution.

Définition 27.40 – Si X et Y sont deux variables aléatoires à valeurs dans E, non nécessairement définies sur le même espace probabilisé, et si $\mathbf{P}_X = \mathbf{P}_Y$, (c'est-à-dire si X et Y suivent la même loi), on note $X \sim Y$.
On dit alors que X et Y sont **identiquement distribuées**.

Remarque. Dire que $X : \Omega_1 \to E$ et $Y : \Omega_2 \to E$ sont identiquement distribuées, c'est dire que $X(\Omega_1) = Y(\Omega_2)$ et que pour tout $x \in X(\Omega_1)$, $\mathbf{P}(X = x) = \mathbf{P}(Y = x)$.

Autrement dit
X et Y prennent les mêmes valeurs, avec les mêmes probas.

27.4.3 Lois usuelles

Si toute famille de réels positifs de somme 1 définit la loi d'une variable aléatoire, certaines, que l'on rencontre fréquemment portent un nom.

La plus simple est aussi la moins intéressante, mais on ne peut l'ignorer :

Définition 27.41 – Une variable aléatoire $X : \Omega \to E$ suit la **loi certaine égale à** $a \in E$ si $a \in X(\Omega)$ et $\mathbf{P}(X = a) = 1$.

On a évidemment envie de voir une loi certaine comme une constante[28], mais une petite distinction peut se produire : on autorise X à prendre d'autres valeurs que a, mais seulement avec probabilité nulle.
Ceci peut notamment se produire lorsqu'un espace probabilisé n'est pas complètement adapté à la description d'une expérience.
Par exemple, on peut modéliser le lancer d'un dé équilibré à 6 faces par $\Omega = [\![1, 7]\!]$, avec $\mathbf{P}(\{i\}) = \begin{cases} \frac{1}{6} & \text{si } 1 \leqslant i \leqslant 6 \\ 0 & \text{si } i = 7 \end{cases}$.
Et on peut alors prendre comme variable aléatoire symbolisant le résultat du premier dé la variable $X : \omega \mapsto \omega$.
Alors $X(7) = 7$, de sorte que $7 \in X(\Omega)$. Mais en revanche, $\mathbf{P}(X = 7) = 0$.
Et alors, la variable aléatoire indicatrice de l'événement «le nombre tiré est strictement inférieur à 7» est une variable certaine égale à 1, mais n'est pas constante égale à 1, car elle vaut 0 pour $\omega = 7$.

[28] Et d'ailleurs les fonctions constantes sur Ω sont des variables aléatoires qui suivent une loi certaine.

Définition 27.42 (Loi de Bernoulli) – Soit $p \in [0, 1]$. Une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathbf{P})$ suit la **loi de Bernoulli de paramètre p** si

$$X(\Omega) = \{0, 1\},\ \mathbf{P}(X = 1) = p,\ \mathbf{P}(X = 0) = 1 - p.$$

Lorsque c'est le cas, on note alors $X \sim \mathscr{B}(p)$.

Remarques. ▶ Une telle variable existe bien car $p + (1 - p) = 1$.
▶ La situation typique de loi de Bernoulli est le PILE (1) ou FACE (0), la pièce tombant sur PILE avec probabilité p. Autrement dit, la variable aléatoire X qui vaut 1 en cas de PILE et 0 en cas de FACE suit la loi $\mathscr{B}(p)$.
Mais plus généralement, toute expérience à deux issues[29] donne lieu à une variable de Bernoulli, 1 désignant l'une des deux issues de l'expérience et 0 l'autre.
▶ Toute variable aléatoire qui ne prend que les valeurs 0 et 1 suit une loi de Bernoulli.
En particulier, la variable indicatrice $\mathbb{1}_A$ d'un événement A suit une loi de Bernoulli. Et puisque $\mathbf{P}(\mathbb{1}_A = 1) = \mathbf{P}(A)$, $\mathbb{1}_A$ suit la loi de Bernoulli de paramètre $\mathbf{P}(A)$.
Inversement, si $X \sim \mathscr{B}(p)$, alors $X = \mathbb{1}_{[X=1]}$.
▶ Enfin, une remarque qui nous servira plus tard : si $X \sim \mathscr{B}(p)$, alors $X^2 = X$.

[29] Qu'on appelle aussi expériences de Bernoulli.

Remarque
Pour une variable qui suit une loi de Bernoulli, la donnée de $\mathbf{P}(X = 1)$ suffit à caractériser entièrement la loi, puisqu'on a alors $\mathbf{P}(X = 0) = 1 - \mathbf{P}(X = 1)$. Ceci vaut plus généralement pour toute variable aléatoire qui ne prend que deux valeurs.

La formule du binôme nous donne une famille de réels positifs dont la somme vaut 1 : si $p \in [0, 1]$, $\displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1$.
Faisons-en donc une loi de variable aléatoire[30] !

[30] C'est le sens de la proposition 27.38 : toute famille de réels positifs de somme 1 permet de définir la loi d'une variable aléatoire.

Définition 27.43 (Loi binomiale) – Soit $p \in]0,1[$, et soit $n \in \mathbf{N}^*$. On dit qu'une variable aléatoire suit la **loi binomiale de paramètres** n **et** p si

$$X(\Omega) = [\![0,n]\!],\ \forall k \in [\![0,n]\!],\ \mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note alors $X \sim \mathscr{B}(n,p)$.

Le contexte où l'on reconnaît naturellement une loi binomiale $\mathscr{B}(n,p)$ est celui de n répétitions **indépendantes** d'épreuves de Bernoulli de probabilité de succès p.
Alors la variable aléatoire qui compte le nombre de succès lors des n répétitions suit la loi binomiale $\mathscr{B}(n,p)$.
Si nous aurons bientôt un argument plus rapide pour prouver ce résultat, expliquons d'où il vient, en notant S_i (resp. E_i) l'événement «la $i^{\text{ème}}$ répétition de l'expérience a amené un succès (resp. un échec)».
Soit alors X la variable qui compte le nombre de succès. Il est clair que $X(\Omega) = [\![0,n]\!]$.
Et on a $[X=0] = E_1 \cap E_2 \cap \cdots \cap E_n$, de sorte que par indépendance des répétitions,

$$\mathbf{P}(X=0) = \mathbf{P}(E_1 \cap E_2 \cap \cdots \cap E_n) = \mathbf{P}(E_1)\mathbf{P}(E_2)\cdots\mathbf{P}(E_n) = (1-p)^n = \binom{n}{0} p^0 (1-p)^{n-0}.$$

Détails
Il s'agit de distinguer n cas suivant la position de l'unique succès.

Ensuite, on a $[X=1] = \displaystyle\bigcup_{i=1}^n \left(S_i \cap \bigcap_{\substack{j=1 \\ j\neq i}}^n E_j \right)$.
Par incompatibilité de ces événements[31], puis par indépendance, il vient

[31] Ceux qui forment l'union.

$$\mathbf{P}(X=1) = \sum_{i=1}^n \mathbf{P}\left(S_i \cap \bigcap_{\substack{j=1 \\ j\neq i}}^n E_j \right) = \sum_{i=1}^n p(1-p)^{n-1} = np(1-p)^{n-1} = \binom{n}{1} p^1 (1-p)^{n-1}.$$

Ensuite, $[X=2] = \displaystyle\bigcup_{1\leqslant i<j\leqslant n} \left(S_i \cap S_j \cap \bigcap_{\substack{k=1 \\ k\neq i,j}}^n E_k \right)$.
Par le même type de raisonnement,

$$\mathbf{P}(X=2) = \sum_{1\leqslant i<j\leqslant n} p^2(1-p)^{n-2} = \binom{n}{2} p^2 (1-p)^{n-2}.$$

En effet, il y a autant de manières de choisir un couple (i,j) avec $1 \leqslant i < j \leqslant n$ qu'il y a de manières de choisir une partie à 2 éléments de $[\![1,n]\!]$, car il n'y aura alors qu'une façon d'ordonner ses éléments. D'où le coefficient $\binom{n}{2}$.
Plus généralement, pour $k \in [\![1,n]\!]$,

$$[X=k] = \bigcup_{\substack{I\subset[\![1,n]\!] \\ \mathrm{Card}(I)=k}} \left(\bigcap_{i\in I} S_i \cap \bigcap_{j\notin I} E_i \right)$$

et comme l'union comporte $\binom{n}{k}$ éléments[32] deux à deux incompatibles, tous de probabilité $p^k(1-p)^{n-k}$, il vient bien

[32] C'est le nombre de k combinaisons de $[\![1,n]\!]$.

$$\mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Concrètement, c'est la loi que vous aurez le plus souvent à reconnaître directement, simplement en remarquant que la variable à laquelle vous avez affaire compte le nombre de succès lors d'une répétition d'épreuves de Bernoulli indépendantes.

☠ Danger !
Méfiez-vous des intuitions que vous pouvez avoir sur l'indépendance...

Exemple 27.44

Un géantiste[33] enfourche à chaque manche avec probabilité p (auquel cas il est disqualifié), les manches étant supposées indépendantes les unes des autres.
Sachant que la saison compte n courses, quelle est la loi de la variable aléatoire X comptabilisant le nombre de courses de deux manches qu'il a terminées ?

Pour chacune des courses, la probabilité qu'il termine les deux manches est $(1-p)^2$.
Et donc X suit la loi binomiale $\mathscr{B}(n,(1-p)^2)$. En revanche, la variable comptant le nombre de courses où le skieur a pris le départ de la seconde manche suit la loi $\mathscr{B}(n, 1-p)$.

[33] Sportif spécialisé dans le slalom géant en ski alpin.

Notons que pour $n=1$, la loi $\mathscr{B}(n,p)$ est la loi de Bernoulli $\mathscr{B}(p)$.

Définition 27.45 (Loi uniforme) – Soit E un ensemble fini. Une variable aléatoire $X : \Omega \to E$ suit la **loi uniforme sur** E si

$$X(\Omega) = E \text{ et } \forall x \in E,\ \mathbf{P}(X=x) = \frac{1}{\mathrm{Card}(E)}.$$

On note alors $X \sim \mathscr{U}(E)$.

Les cas particuliers qui nous intéresserons seront ceux où $E = [\![a,b]\!]$, où $a<b$ sont deux entiers relatifs. Alors si $X \sim \mathscr{U}([\![a,b]\!])$, pour tout $k \in [\![a,b]\!]$, $\mathbf{P}(X=k) = \dfrac{1}{b-a+1}$.
En particulier, si $X \sim \mathscr{U}([\![1,n]\!])$, alors $\forall k \in [\![1,n]\!]$, $\mathbf{P}(X=k) = \dfrac{1}{n}$.

Ces lois se rencontrent essentiellement lors de lancers de dés équilibrés/ de tirages dans une urne, lorsqu'on est en situation d'équiprobabilité.