31 Probabilités sur un univers fini, variables aléatoires et lois

1 Événements/variables aléatoires associés à une expérience aléatoire

Le concept d'*expérience aléatoire* n'est pas mathématique à proprement parler. On appelle ainsi toute expérience susceptible *a priori* de *résultats* différents lors de répétitions dans des conditions similaires. L'ensemble de ces résultats observables est appelé l'*univers* de l'expérience étudiée et est souvent noté Ω. Plutôt que de résultats, on parle aussi souvent d'*issues* ou de *réalisations*.

- *Lancer d'un dé à 6 faces.* L'expérience d'un lancer de dé à 6 faces peut conduire à 6 résultats, selon la face obtenue. Pour modéliser un tel lancer, on peut donc choisir pour univers Ω l'ensemble $[\![1\,,6]\!]$.
- *Deux lancers d'un dé à 6 faces.* Si l'on lance deux fois de suite un même dé, la modélisation la plus naturelle des résultats est offerte par l'univers $\Omega = [\![1\,,6]\!]^2$, où la première coordonnée d'un couple $(i,j) \in \Omega$ représente le résultat du premier lancer et la seconde celui du second lancer. Plus généralement, n lancers d'un même dé seront modélisés par l'univers $[\![1\,,6]\!]^n$.
- *Lancer de deux dés.* Lançons maintenant deux dés identiques une seule fois. On peut ici aussi modéliser les issues par des couples (i,j) de nombres entre 1 et 6. Il faut alors veiller à ce que les couples $(2,4)$ et $(4,2)$ représentent la même issue. Ainsi l'univers sera $\Omega = \left\{(i,j) \in [\![1\,,6]\!]^2 \;\middle|\; i \leqslant j\right\}$. Ceci dit, cette expérience n'est modifiée que marginalement si l'on suppose l'un des deux dés peint en rouge et l'autre en vert. L'univers s'écrit alors à nouveau $\Omega = [\![1\,,6]\!]^2$, en convenant par exemple que la première coordonnée représente le résultat du dé rouge et la seconde celui du dé vert.

Ainsi une même expérience aléatoire peut être décrite de diverses façons et sa modélisation dépend donc d'un choix du probabiliste. Deux descriptions concurrentes ne conduisent pas nécessairement à la même étude, mais doivent donner des résultats observables identiques. On choisira alors la plus commode.

- *Lancer d'une pièce de monnaie.* L'expérience d'un lancer d'une pièce de monnaie peut conduire à 2 issues, selon que l'on obtient Pile ou Face, on peut donc choisir l'univers $\{\text{Pile}, \text{Face}\}$. On peut aussi, afin d'alléger les notations, utiliser l'univers $\{0,1\}$, où l'on convient que 0 représente l'obtention d'un Pile et 1 celui d'un Face.
- *Lancers répétés d'une pièce de monnaie.* On lance maintenant une pièce de monnaie jusqu'à obtenir un Face. Avec la convention de notation précédente, on peut alors considérer l'univers $\{\infty, 1, 01, 001, 0001, \ldots\}$, qui contient tous les mots écrits avec une suite quelconque de 0 terminée par un 1 et où le symbole ∞ représente l'issue « La pièce donne toujours Pile ». Mais on pourrait aussi utiliser l'univers $\Omega = \{0,1,2,3,4,\ldots\} = \mathbb{N}$, où chaque entier désigne le rang d'apparition de Face. Il y a une bijection évidente entre ces deux ensembles et l'univers de cette expérience est infini en bijection avec $\mathbb{N}$.
- *Lancers répétés d'une pièce de monnaie (bis).* On pourrait aussi vouloir observer des suites quelconques de Pile et de Face. L'univers attaché naturellement à ces situations est $\{0,1\}^{\mathbb{N}}$, l'ensemble des suites infinies de 0 et de 1.

Dans ce chapitre, tous les univers considérés seront finis. Certains univers infinis seront étudiés en deuxième année.

Pour une expérience aléatoire donnée, on peut bien sûr appréhender chaque résultat isolément en tant qu'ÉLÉMENT de Ω, mais généralement notre intérêt s'oriente vers des regroupements d'issues définies génériquement par une propriété commune et appelés *événements*.

- Dans le lancer de dé précédent, la propriété « La face obtenue est paire » est satisfaite par trois résultats, en l'occurrence 2, 4 et 6, et sera identifiée à l'ensemble $\{2,4,6\}$, *i.e.* une partie de Ω.
- Lors du lancer de deux dés, la propriété « la somme des deux dés obtenus vaut 4 » est, dans le cadre de la modélisation par l'univers $\Omega = [\![1\,,6]\!]^2$ (avec deux dés pouvant être distingués), satisfaite par les issues $(1,3)$, $(2,2)$ et $(3,1)$ et sera identifiée à l'ensemble $\{(1,3),(2,2),(3,1)\}$, *i.e.* une partie de Ω.
- Lors de lancers répétés d'une pièce, on peut s'intéresser aux événements « Obtenir la séquence 010 avant la séquence 111 » ou encore « Ne jamais observer deux 1 consécutifs », etc..

Ainsi les événements sont des PARTIES de l'univers Ω qui modélise l'expérience aléatoire considérée.

Définition 1 – Vocabulaire usuel sur les événements

Soit Ω l'univers fini d'une expérience aléatoire, *i.e.* l'ensemble des issues possibles de cette expérience.

- On appelle *événement* toute partie de Ω, *i.e.* tout élément de $\mathscr{P}(\Omega)$. Un événement A est dit *réalisé* lorsque le résultat de l'expérience est un élément de A.
- Les singletons $\{\omega\}$, avec $\omega \in \Omega$, sont appelés les *événements élémentaires* de Ω.
- L'événement Ω est l'*événement certain* et l'événement $\varnothing$ est l'*événement impossible*.
- Pour tous événements $A, B \in \mathscr{P}(\Omega)$,
  - × l'événement $A \cup B$, appelé l'événement « A ou B », est réalisé lorsque l'un au moins des événements A et B l'est ;
  - × l'événement $A \cap B$, appelé l'événement « A et B », est réalisé lorsque les événements A et B le sont ;
  - × l'événement $\overline{A}$, appelé l'*événement contraire à* A, est réalisé lorsque A ne l'est pas.
- Deux événements A et B de Ω sont dits *incompatibles* lorsqu'ils sont disjoints, *i.e.* $A \cap B = \varnothing$.
- On dit enfin que l'événement A *implique* l'événement B lorsque $A \subset B$.

En termes probabilistes, pour toute famille $(A_1, \ldots, A_n)$ d'événements d'un univers Ω,

- l'événement $\bigcup\limits_{1\leqslant i\leqslant n} A_i$ correspond à « au moins l'un des événements de la famille $(A_1, \ldots, A_n)$ est réalisé » ;
- l'événement $\bigcap\limits_{1\leqslant i\leqslant n} A_i$ correspond à « tous les événements de la famille $(A_1, \ldots, A_n)$ sont réalisés ».

Exemple 2 Observons par exemple n personnes passant la porte d'un bureau de poste. On note, pour $k \in [\![1\,, n]\!]$, F_k l'événement « la k^{e} personne est une femme ». Il est alors aisé d'écrire formellement les événements suivants :

- « les n personnes sont des femmes » est traduit par $\bigcap\limits_{1\leqslant k\leqslant n} F_k$;
- « au moins une des n personnes est une femme » est traduit par $\bigcup\limits_{1\leqslant k\leqslant n} F_k$;
- « aucune des n personnes n'est une femme » est traduit par $\overline{\bigcup\limits_{1\leqslant k\leqslant n} F_k} = \bigcap\limits_{1\leqslant k\leqslant n} \overline{F}_k$.

Définition 3 – Système complet d'événements

Soit Ω un univers fini. On appelle *système complet d'événements de* Ω toute famille finie $(A_i)_{i\in I}$, où I un ensemble fini, d'événements deux à deux incompatibles et dont la réunion est l'événement certain, autrement dit pour laquelle $\Omega = \bigsqcup\limits_{i\in I} A_i$ – réunion disjointe.

Utiliser un système complet d'événements revient à raisonner par disjonction des cas. On découpe l'univers en plusieurs événements, un et un seul d'entre eux étant réalisé lors de chaque expérience.

Exemple 4

- Pour l'expérience aléatoire du lancer d'un dé à 6 faces, les événements « La face obtenue est paire » et « La face obtenue est impaire » forment un système complet d'événements. En effet, on obtient à tous les coups une face paire ou impaire, mais jamais les deux simultanément.
- Plus généralement, pour tout événement A d'un univers fini Ω, le couple d'événements $(A, \overline{A})$ est un système complet d'événements, puisque $A \cap \overline{A} = \varnothing$ et $A \cup \overline{A} = \Omega$.
- Si l'univers est un jeu de 52 cartes, on peut découper celui-ci en divers systèmes complets d'événements. Par exemple, les quatre événements $A_\clubsuit$ « la carte est un trèfle », $A_\spadesuit$, $A_\heartsuit$ et $A_\diamondsuit$ forment un système complet d'événements. Mais on pourrait aussi considérer celui formé par les treize événements B_i « la carte est un i », avec $i \in \{1, 2, \ldots, 10, V, D, R\}$.
- Pour tout univers fini $\Omega = \{\omega_1, \ldots, \omega_n\}$, il est clair que $\Omega = \bigsqcup\limits_{i=1}^{n} \{\omega_i\}$, ainsi la famille des événements élémentaires $(\{\omega_1\}, \ldots, \{\omega_n\})$ forme un système complet d'événements.

La plupart du temps, les événements que l'on manipule sont construits à partir de grandeurs, numériques ou non, appelées des variables aléatoires. Par exemple, quand on lance une pièce 10 fois successivement, l'événement « On a obtenu exactement 3 piles » peut être écrit « $N = 3$ » si l'on note N le nombre de piles obtenus. Ce nombre N dépend des lancers que l'on effectue, *i.e.* des réalisations liées à notre expérience aléatoire, et doit donc être vu comme une fonction.

Définition 5 – Variable aléatoire (réelle), événements associés

Soit Ω un univers fini et E un ensemble.

- On appelle *variable aléatoire sur Ω à valeurs dans E* toute application X de Ω dans E.
  Une telle variable aléatoire est parfois qualifiée de *finie* dans la mesure où l'ensemble $X[\Omega]$ des valeurs prises par X est fini à l'instar de Ω.
  Lorsque $E \subset \mathbb{R}$, on parle de *variable aléatoire réelle.*
- Pour toute partie A de E, on note $(X \in A)$ ou $\{X \in A\}$ l'événement $X^{-1}[A] = \{\omega \in \Omega \mid X(\omega) \in A\}$.
  Lorsque $A = \{a\}$ est un singleton, on privilégie la notation $(X = a)$ ou $\{X = a\}$ pour désigner l'événement
  $$X^{-1}[\{a\}] = \{\omega \in \Omega \mid X(\omega) = a\}.$$
  Lorsque $E = \mathbb{R}$ et $A =]-\infty\,, a]$ est un intervalle, on privilégie la notation $(X \leqslant a)$ ou $\{X \leqslant a\}$ pour désigner l'événement
  $$X^{-1}\big[]-\infty\,, a]\big] = \{\omega \in \Omega \mid X(\omega) \leqslant a\}.$$
  On définit évidemment de même $(X \geqslant a)$, $(X < a)$ et $(X > a)$.

Une variable aléatoire X associe une valeur à chaque issue lié à une expérience aléatoire. Mais au-delà de cette transformation, elle redéfinit complètement l'espace d'étude : tout sous-ensemble de E devient au travers de X un événement « étudiable ».

Rappelons que l'opération « image réciproque » est compatible avec les opérations ensemblistes :

$$X^{-1}[\varnothing] = \varnothing \qquad\qquad X^{-1}\big[\,\overline{A}\,\big] = \overline{X^{-1}[A]}$$
$$X^{-1}[A \cup B] = X^{-1}[A] \cup X^{-1}[B] \qquad\qquad X^{-1}[A \cap B] = X^{-1}[A] \cap X^{-1}[B],$$

où A et B sont deux parties de E. On peut alors manipuler naturellement les notations introduites précédemment.

Exemple 6 Si X est une variable aléatoire réelle,

$$(X \leqslant 1) \cap (X \geqslant 0) = X^{-1}\big[]-\infty\,, 1]\big] \cap X^{-1}\big[[0\,, +\infty[\big] = X^{-1}\big[]-\infty\,, 1] \cap [0\,, +\infty[\big] = X^{-1}\big[[0\,, 1]\big] = (0 \leqslant X \leqslant 1).$$

De façon générale, pour deux parties A et B de E,

$$(X \in A \cup B) = (X \in A) \cup (X \in B) \qquad \text{et} \qquad (X \in A \cap B) = (X \in A) \cap (X \in B).$$

Remarque 7 L'appellation variable aléatoire, bien que malheureuse, est usuelle. En effet, X n'est pas une variable, mais bien une fonction et celle-ci n'a rien d'aléatoire, mais est plutôt entièrement déterministe (à toute issue ω de l'expérience aléatoire, X associe une valeur dans E totalement déterminée). L'aléatoire fera irruption quand nous attribuerons à tout événement de Ω une probabilité, *i.e.* une certaine mesure de sa vraisemblance.

Exemple 8 – Variable aléatoire certaine Soit $c \in E$. L'application constante $X : \omega \longmapsto c$ de Ω dans E est une variable aléatoire, appelée *variable aléatoire certaine* égale à c.

Exemple 9 Pour modéliser le lancer d'un dé à 6 faces 2 fois, on peut choisir pour univers Ω l'ensemble $[\![1\,, 6]\!]^2$ des 2-listes de $[\![1\,, 6]\!]$. La fonction X_1 (resp. X_2) « valeur obtenue au premier (resp. deuxième) lancer » est une variable aléatoire réelle finie sur Ω d'image $X_1[\Omega] = [\![1\,, 6]\!]$. La fonction $S = X_1 + X_2$ « somme des deux faces obtenues » en est une autre, d'image $[\![2\,, 12]\!]$.
Par exemple, pour l'issue $\omega = (2, 5)$, $X_1(\omega) = 2$, $X_2(\omega) = 5$ et $S(\omega) = 7$.

Exemple 10 Un concierge possède un trousseau de 10 clés, dont une seule permet d'ouvrir la porte face à lui. On suppose que les essaies de clés se font au hasard et SANS remise et on note X le nombre de clés essayées pour ouvrir la porte. Alors $X[\Omega] = [\![1\,, 10]\!]$ et X est une variable aléatoire réelle finie.

Définition-théorème 11 – Système complet d'événements associé à une variable aléatoire

Soit X une variable aléatoire sur un univers fini Ω. La famille d'événements $((X = x))_{x \in X[\Omega]}$ forme un système complet d'événements de l'univers Ω, appelée *système complet d'événements associé à X*. En outre,

$$\forall A \in \mathscr{P}(X[\Omega]), \quad (X \in A) = \bigsqcup_{x \in A} (X = x).$$

Démonstration. On a tout simplement $\Omega = \bigsqcup_{x \in X[\Omega]} (X = x)$. ■

Exemple 12 Le système complet d'événements associé à la variable aléatoire S de l'exemple 9 est : $((S = k))_{2 \leqslant k \leqslant 12}$.

2 Espaces probabilisés finis

Le concept d'expérience aléatoire décrit des situations susceptibles de conduire à différents résultats, mais rien ne nous permet pour le moment de mesurer la vraissemblance de ces résultats les uns par rapport aux autres. Nous allons associer dans ce paragraphe à tout événement d'une expérience aléatoire une probabilité, *i.e.* un réel compris entre 0 et 1 pour lequel la valeur 0 représente le plus bas niveau de vraisemblance et la valeur 1 le niveau le plus élevé.

2.1 Définition et propriétés

Définition 13 – Probabilité sur un univers fini, espace probabilisé fini

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

(i) $\mathrm{P}(\Omega) = 1$;

(ii) $\forall A, B \in \mathscr{P}(\Omega), \quad A \cap B = \varnothing \implies \mathrm{P}(A \sqcup B) = \mathrm{P}(A) + \mathrm{P}(B)$ *(additivité finie)*.

Le couple (Ω, P) ainsi formé est appelé un *espace probabilisé fini*.

Remarque 14 Une probabilité est par définition à valeurs dans l'intervalle $[0\,,1]$, ainsi, pour tout événement $A \in \mathscr{P}(\Omega)$, on a $0 \leqslant \mathrm{P}(A) \leqslant 1$. On vérifiera donc SYSTÉMATIQUEMENT que le résultat d'un calcul de probabilité est un nombre compris entre 0 et 1.

Théorème 15 – Propriétés des probabilités

Soit (Ω, P) un espace probabilisé fini et $A, B, A_1, \ldots, A_n \in \mathscr{P}(\Omega)$.

(i) Ensemble vide. $\mathrm{P}(\varnothing) = 0$.

(ii) Complémentaire et différence. $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$ et $\mathrm{P}(A \backslash B) = \mathrm{P}(A) - \mathrm{P}(A \cap B)$.

(iii) Croissance. Si $A \subset B$, alors $\mathrm{P}(A) \leqslant \mathrm{P}(B)$.

(iv) Union quelconque. $\mathrm{P}(A_1 \cup A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \cap A_2)$.

(v) Union disjointe. Si $A_1, \ldots, A_n$ sont DEUX À DEUX INCOMPATIBLES, alors $\mathrm{P}\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathrm{P}(A_k)$.

Démonstration.

(ii) Comme $A = (A \backslash B) \sqcup (A \cap B)$, $\mathrm{P}(A) = \mathrm{P}(A \backslash B) + \mathrm{P}(A \cap B)$.
En particulier, $\mathrm{P}(\overline{A}) = \mathrm{P}(\Omega \backslash A) = \mathrm{P}(\Omega) - \mathrm{P}(\Omega \cap A) = 1 - \mathrm{P}(A)$.

(i) $\mathrm{P}(\varnothing) = 1 - \mathrm{P}(\Omega) = 1 - 1 = 0$.

(iii) Puisque $A \subset B$, $B = A \sqcup (B \backslash A)$. Or $\mathrm{P}(B \backslash A) \geqslant 0$, ainsi $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(B \backslash A) \geqslant \mathrm{P}(A)$.

(iv) D'une part, $\mathrm{P}(A) = \mathrm{P}(A \backslash B) + \mathrm{P}(A \cap B)$ (cf. deuxième point), et d'autre part $\mathrm{P}(A \cup B) = \mathrm{P}(A \backslash B) + \mathrm{P}(B)$, puisque $A \cup B = (A \backslash B) \sqcup B$. Il suffit alors de mélanger...

(v) On procède par récurrence sur $n \in \mathbb{N}^*$, via l'additivité de P.

■

Remarque 16 On pourra noter l'analogie de ces formules avec celles concernant les calculs de cardinaux (théorème 5 du chapitre 30). La formule du crible du point **(iv)** se généralise notamment à un entier n quelconque (hors programme, cf. chapitre 32 pour une preuve) :

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \mathrm{P}(A_{i_1} \cap \ldots \cap A_{i_k}).$$

Corollaire 17

Si $(A_i)_{i \in I}$ est un système complet d'événements d'un espace probabilisé fini (Ω, P), alors $\displaystyle\sum_{i \in I} \mathrm{P}(A_i) = 1$.

Démonstration. Par définition d'un système complet d'événements, $\Omega = \displaystyle\bigsqcup_{i \in I} A_i$. ∎

Vue comme une application entre l'ensemble des événements $\mathscr{P}(\Omega)$ et l'intervalle $[0\,,1]$, une probabilité est un objet *a priori* complexe à définir. Pour une expérience aussi simple que le lancer d'un dé, doit-on expliciter la probabilité des 2^6 événements de $\mathscr{P}([\![1\,,6]\!])$? Puis vérifier les axiomes de la définition ? Un autre angle d'attaque est nettement préférable et consiste à se limiter à une classe particulière d'événements dont on définit la probabilité, puis à étendre cette dernière aux autres événements, de façon à respecter les axiomes d'une probabilité. Ce principe est l'objet du théorème suivant.

Définition 18 – Distribution de probabilités sur un ensemble fini

Soit I un ensemble fini. Une *distribution de probabilités sur I* est une famille $(p_i)_{i \in I}$ de réels POSITIFS indexée par I et de somme 1.

Exemple 19 Soit (Ω, P) un espace probabilisé fini. La famille $(\mathrm{P}(\{\omega\}))_{\omega \in \Omega}$ des probabilités des événements élémentaires de l'univers Ω est une distribution de probabilités sur Ω.

En effet, $\displaystyle\sum_{\omega \in \Omega} \mathrm{P}(\{\omega\}) = \mathrm{P}\left(\bigsqcup_{\omega \in \Omega} \{\omega\}\right) = \mathrm{P}(\Omega) = 1$.

Théorème 20 – Détermination d'une probabilité sur les événements élémentaires

Une probabilité sur un univers Ω fini est entièrement déterminée par ses valeurs sur les événements élémentaires, *i.e.* par la distribution de probabilités $(\mathrm{P}(\{\omega\}))_{\omega \in \Omega}$.

Précisément, si l'univers $\Omega = \{\omega_1, \ldots, \omega_n\}$ est fini et si $(p_i)_{1 \leqslant i \leqslant n}$ est une famille de réels positifs de somme 1, alors il existe une unique probabilité P sur Ω telle que, pour tout $i \in [\![1\,,n]\!]$, $\mathrm{P}(\{\omega_i\}) = p_i$.

Démonstration.

- *Analyse.* Soit P une probabilité sur Ω telle que, pour tout $i \in [\![1\,,n]\!]$, $\mathrm{P}(\{\omega_i\}) = p_i$. Alors, pour tout $A \in \mathscr{P}(\Omega)$,

$$\mathrm{P}(A) = \mathrm{P}\left(\bigsqcup_{\omega_i \in A} \{\omega_i\}\right) = \sum_{\omega_i \in A} \mathrm{P}(\{\omega_i\}) = \sum_{\omega_i \in A} p_i = \sum_{i=1}^{n} \mathbb{1}_A(\omega_i) p_i.$$

  En particulier, la probabilité P est unique sous réserve d'existence.

- *Synthèse.* Pour tout $A \in \mathscr{P}(\Omega)$, posons $\mathrm{P}(A) = \displaystyle\sum_{i=1}^{n} \mathbb{1}_A(\omega_i) p_i$.

  × Pour tout $A \in \mathscr{P}(\Omega)$, $\mathrm{P}(A) \in [0\,,1]$, car $0 \leqslant \mathrm{P}(A) \leqslant \displaystyle\sum_{i=1}^{n} p_i = 1$.

  × $\mathrm{P}(\Omega) = \displaystyle\sum_{i=1}^{n} \mathbb{1}_\Omega(\omega_i) p_i = \sum_{i=1}^{n} p_i = 1$.

  × Pour tout $j \in [\![1\,,n]\!]$, $\mathrm{P}(\{\omega_j\}) = \displaystyle\sum_{i=1}^{n} \mathbb{1}_{\{\omega_j\}}(\omega_i) p_i = \sum_{i=1}^{n} \delta_{i,j} p_i = p_j$.

  × Pour tous $A, B \in \mathscr{P}(\Omega)$ incompatibles, sachant $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B} = \mathbb{1}_A + \mathbb{1}_B$,

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

∎

Exemple 21 – Loi binomiale Considérons l'univers $\Omega = [\![0\,,n]\!]$, un réel $p \in]0\,,1[$ et posons, pour tout $k \in [\![0\,,n]\!]$,

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}.$$

Cette famille de réels définie une probabilité sur Ω.

En effet, ces réels sont positifs et $\displaystyle\sum_{k=0}^{n} p_k = (p+1-p)^n = 1$, d'après la formule du binôme.

Exemple 22 Une urne contient $3n$ boules numérotées de 1 à $2n$ avec, pour tout $k \in [\![1\,,n]\!]$, exactement 2 boules indiscernables de numéro $2k$ et exactement une boule de numéro $2k-1$. On tire de cette urne une boule au hasard et on regarde son numéro. Pour modéliser ce tirage, on peut choisir pour univers Ω l'ensemble $[\![1\,,2n]\!]$ de tous les numéros possibles et pour probabilité P la probabilité définie, pour tout $k \in [\![1\,,n]\!]$, par $\mathrm{P}(\{2k\}) = \dfrac{2}{3n}$ et $\mathrm{P}(\{2k-1\}) = \dfrac{1}{3n}$.

En effet, on définit bien ainsi une probabilité, puisque la somme des probabilités élémentaires, qui sont des réels positifs, est égale à 1 :

$$\sum_{k=1}^{2n} \mathrm{P}(\{k\}) = \sum_{k=1}^{n} \mathrm{P}(\{2k\}) + \sum_{k=1}^{n} \mathrm{P}(\{2k-1\}) = \sum_{k=1}^{n} \frac{2}{3n} + \sum_{k=1}^{n} \frac{1}{3n} = 1.$$

L'événement A « On tire une boule de numéro pair » est alors de probabilité $\displaystyle\mathrm{P}(A) = \sum_{k=1}^{n} \mathrm{P}(\{2k\}) = \sum_{k=1}^{n} \frac{2}{3n} = \frac{2}{3}$.

2.2 Un exemple fondamental : la probabilité uniforme

Définition-théorème 23 – Événements équiprobables, probabilité uniforme

Soit (Ω, P) un espace probabilisé fini.

- Deux événements A et B sont dit *équiprobables* lorsque $\mathrm{P}(A) = \mathrm{P}(B)$.
- La probabilité P est dite *uniforme* lorsque tous les événements élémentaires sont équiprobables, ce qui revient à avoir $\mathrm{P}(\{\omega\}) = \dfrac{1}{|\Omega|}$, pour tout $\omega \in \Omega$. Le cas échéant, pour tout $A \in \mathscr{P}(\Omega)$, $\mathrm{P}(A) = \dfrac{|A|}{|\Omega|}$.

Démonstration. L'existence d'une telle probabilité est assurée par le théorème 20. ∎

La relation « $\mathrm{P}(A) = \dfrac{|A|}{|\Omega|}$ » est naturelle et bien connue, c'est le fameux rapport « $\dfrac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}}$ », qui ramène dans le contexte de la probabilité uniforme tout calcul de probabilité à un problème de dénombrement. L'hypothèse d'équiprobabilité s'applique lorsqu'aucun événement élémentaire n'est favorisé par rapport à un autre (dés non pipés, pièces équilibrées, boules indiscernables dans une urne, ...).

Exemple 24 Pour la variable aléatoire S de l'exemple 9, en munissant l'univers $\Omega = [\![1\,,6]\!]^2$ de la probabilité uniforme,

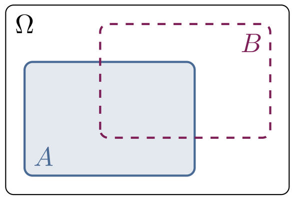
- $(S=8) = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$ et ainsi $\mathrm{P}(S=8) = \dfrac{5}{36}$;
- $(S \leqslant 4) = \{(1,1),(1,2),(2,1),(1,3),(2,2),(3,1)\}$ et ainsi $\mathrm{P}(S \leqslant 4) = \dfrac{6}{36} = \dfrac{1}{6}$.

3 Probabilité conditionnelle

En général, lorsqu'un processus expérimental complexe est mis en jeu, le résultat final est difficile à décrire directement. Il est alors préférable de décomposer l'expérience globale en plusieurs étapes. La notion de probabilité conditionnelle permet d'étudier séparément chacune de ces étapes.

3.1 Définition

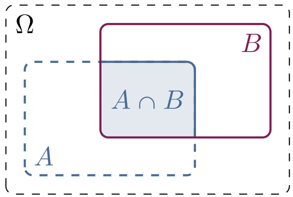
Soit Ω un univers fini, $A, B \in \mathscr{P}(\Omega)$ et notons P la probabilité UNIFORME sur Ω. Dans ce cadre, l'événement A a pour probabilité $\mathrm{P}(A) = \dfrac{|A|}{|\Omega|}$.



Faisons maintenant l'hypothèse que $\mathrm{P}(B) > 0$ et que nous disposons de l'information supplémentaire suivante concernant la situation modélisée par (Ω, P) : l'événement B est réalisé. La probabilité P choisie initialement n'est alors plus pertinente puisque, B étant réalisé, on devrait avoir $\mathrm{P}(B) = 1$ et $\mathrm{P}(\overline{B}) = 0$, ce qui n'est pas le cas *a priori*. Il convient donc d'introduire une nouvelle probabilité P_B, mais selon quelle modalité, sachant que B est réalisé ? Il est possible de conserver l'univers Ω, tout en considérant qu'il n'est plus l'ensemble de tous les cas possibles de l'expérience modélisée, qui sont dorénavant circonscrits à B. Quant à l'ensemble des cas favorables à la réalisation de A, celui-ci n'est plus A tout entier, mais sa restriction $A \cap B$.

Finalement, la nouvelle probabilité $\mathrm{P}_B(A)$ de A vaut :

$$\mathrm{P}_B(A) = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$



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

Relativement à P, la probabilité $\mathrm{P}_B : A \longmapsto \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$, dite *probabilité conditionnelle sachant* B, se présente comme une fonction d'oubli, une émanation de P aveugle à tout ce qui dépasse le cadre de B. La division par $\mathrm{P}(B)$ garantit l'égalité $\mathrm{P}_B(B) = 1$, selon laquelle « B est tout » – un nouvel univers apparent. En résumé :

Conditionner revient à redimensionner l'ensemble des cas possibles.

De façon générale, on adopte la définition suivante.

Définition-théorème 25 – Probabilité conditionnelle

Soit (Ω, P) un espace probabilisé fini et $B \in \mathscr{P}(\Omega)$ un événement de probabilité NON NULLE.

- Pour tout $A \in \mathscr{P}(\Omega)$, la *probabilité conditionnelle de A sachant B*, notée $\mathrm{P}_B(A)$ ou $\mathrm{P}(A|B)$, est le réel

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

- L'application $\mathrm{P}_B : \mathscr{P}(\Omega) \longrightarrow [0\,,1]$ est une probabilité sur Ω, appelée la *probabilité conditionnée à B*. En particulier, l'application P_B jouit des propriétés d'une probabilité – listées au théorème 15.

Démonstration. Il s'agit de montrer que P_B est une probabilité sur Ω.

- Pour tout $A \in \mathscr{P}(\Omega)$, $\mathrm{P}_B(A) \in [0\,,1]$, car $0 \leqslant \mathrm{P}(A \cap B) \leqslant \mathrm{P}(B) \leqslant 1$, par croissance de P.
- En outre, $\mathrm{P}_B(\Omega) = \dfrac{\mathrm{P}(\Omega \cap B)}{\mathrm{P}(B)} = \dfrac{\mathrm{P}(B)}{\mathrm{P}(B)} = 1$.
- Enfin, pour tous $A, A' \in \mathscr{P}(\Omega)$ incompatibles, P étant additive,

$$\mathrm{P}_B(A \sqcup A') = \frac{\mathrm{P}\big((A \sqcup A') \cap B\big)}{\mathrm{P}(B)} = \frac{\mathrm{P}\big((A \cap B) \sqcup (A' \cap B)\big)}{\mathrm{P}(B)} = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} + \frac{\mathrm{P}(A' \cap B)}{\mathrm{P}(B)} = \mathrm{P}_B(A) + \mathrm{P}_B(A').$$

ainsi P_B est additive.

■

La probabilité de A sachant B est la probabilité de l'événement A dans la situation où l'on sait que l'événement B s'est réalisé.

Remarque 26 On ne sait pas définir $\mathrm{P}_B(A)$ lorsque $\mathrm{P}(B) = 0$. Toutefois, par convention, on considère que l'égalité

$$0 = \mathrm{P}(A \cap B) = \mathrm{P}(B)\mathrm{P}_B(A)$$

reste valable dans ce cas. De fait, si $\mathrm{P}(B) = 0$, alors par croissance de P : $0 \leqslant \mathrm{P}(A \cap B) \leqslant \mathrm{P}(B) = 0$.

Exemple 27 On lance un dé cubique équilibré. L'univers est donc $[\![1\,,6]\!]$ et est muni de la probabilité uniforme, notée P. On considère les événements suivants :

- A : « obtenir un nombre inférieur ou égal à 3 » ;
- B : « obtenir un 5 » ;
- C : « obtenir un 2 ».

Calculons $\mathrm{P}_A(B)$ et $\mathrm{P}_A(C)$.

- Il y a deux façons d'envisager les choses pour $\mathrm{P}_A(B)$:
  1. Par définition :
  $$\mathrm{P}_A(B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} = \frac{\mathrm{P}(\varnothing)}{\mathrm{P}(A)} = 0.$$
  2. $\mathrm{P}_A(B)$ est la probabilité que B soit réalisé sachant que A l'est. Or, sachant que le résultat du lancer est inférieur à 3, il est impossible qu'il soit égal à 5.
- Il y a deux façons d'envisager les choses pour $\mathrm{P}_A(C)$:
  1. Par définition :
  $$\mathrm{P}_A(C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(A)} = \frac{\mathrm{P}(C)}{\mathrm{P}(A)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$
  2. Si A est réalisé, l'issue du lancer est un élément de $\{1, 2, 3\}$. Les résultats étant équiprobables, la probabilité d'avoir obtenu un 2 dans cet univers des possibles est alors de $\dfrac{1}{|\{1, 2, 3\}|} = \dfrac{1}{3}$.

3.2 Formules en lien avec les probabilités conditionnelles

Dans cette section, (Ω, P) désigne un espace probabilisé fini.

3.2.1 Formule des probabilités composées

La formule des probabilités composées généralise la relation $\mathrm{P}(A \cap B) = \mathrm{P}(B)\mathrm{P}_B(A)$ à une intersection finie quelconque d'événements.

Théorème 28 – Formule des probabilités composées

Si $A_1, \ldots, A_n \in \mathscr{P}(\Omega)$, alors

$$\mathrm{P}(A_1 \cap \ldots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}_{A_1}(A_2)\mathrm{P}_{A_1 \cap A_2}(A_3) \ldots \mathrm{P}_{A_1 \cap \ldots \cap A_{n-1}}(A_n) = \prod_{i=1}^{n} \mathrm{P}_{\bigcap_{j=1}^{i-1} A_j}(A_i).$$

Démonstration. Si $\mathrm{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$, alors, par croissance de P, $\mathrm{P}(A_1 \cap \ldots \cap A_i) > 0$, pour tout $1 \leqslant i \leqslant n-1$. Ainsi toutes les probabilités conditionnelles en jeu sont licites et une simplification télescopique permet de conclure

$$\prod_{i=1}^{n} \mathrm{P}_{\bigcap_{j=1}^{i-1} A_j}(A_i) = \mathrm{P}(A_1) \prod_{i=2}^{n} \frac{\mathrm{P}(A_1 \cap \ldots \cap A_i)}{\mathrm{P}(A_1 \cap \ldots \cap A_{i-1})} = \mathrm{P}(A_1 \cap \ldots \cap A_n).$$

Sinon il existe un plus petit entier $i \in [\![1\,, n-1]\!]$ tel que $\mathrm{P}(A_1 \cap \ldots \cap A_i) = 0$. Par croissance de la probabilité P, il vient d'une part $\mathrm{P}(A_1 \cap \ldots \cap A_n) = 0$. D'autre part, $\mathrm{P}(A_1 \cap \ldots \cap A_{i-1}) > 0$ et, d'après le point précédent,

$$\mathrm{P}(A_1)\mathrm{P}_{A_1}(A_2)\mathrm{P}_{A_1 \cap A_2}(A_3) \ldots \mathrm{P}_{A_1 \cap \ldots \cap A_{n-1}}(A_n) = \mathrm{P}(A_1 \cap \ldots \cap A_i)\mathrm{P}_{A_1 \cap \ldots \cap A_i}(A_{i+1}) \ldots \mathrm{P}_{A_1 \cap \ldots \cap A_{n-1}}(A_n) = 0,$$

par convention. ∎

Cette formule s'utilise dès qu'il est possible de décomposer l'expérience aléatoire en plusieurs étapes successives.

Exemple 29 Une urne contient 5 boules blanches et 8 boules noires indiscernables. L'expérience consiste à tirer successivement et sans remise 3 boules. Quelle est la probabilité que les trois boules tirées soient blanches ?

Exemple 30 Un commerçant met en vente n tickets d'un certain jeu dont exactement g sont gagnants. Je lui achète k tickets, avec $k \leqslant n - g$. Avec quelle probabilité en ai-je acheté au moins un gagnant ?

3.2.2 Formule des probabilités totales

Théorème 31 – Formule des probabilités totales

Si $(A_1, \dots, A_n)$ est un système complet d'événements, alors

$$\forall B \in \mathscr{P}(\Omega), \quad \mathrm{P}(B) = \sum_{i=1}^{n} \mathrm{P}(A_i \cap B) = \sum_{i=1}^{n} \mathrm{P}(A_i)\mathrm{P}_{A_i}(B).$$

Démonstration. Soit $B \in \mathscr{P}(\Omega)$, alors

$$B = B \cap \Omega = B \cap \left(\bigsqcup_{i=1}^{n} A_i \right) = \bigsqcup_{i=1}^{n} (B \cap A_i).$$

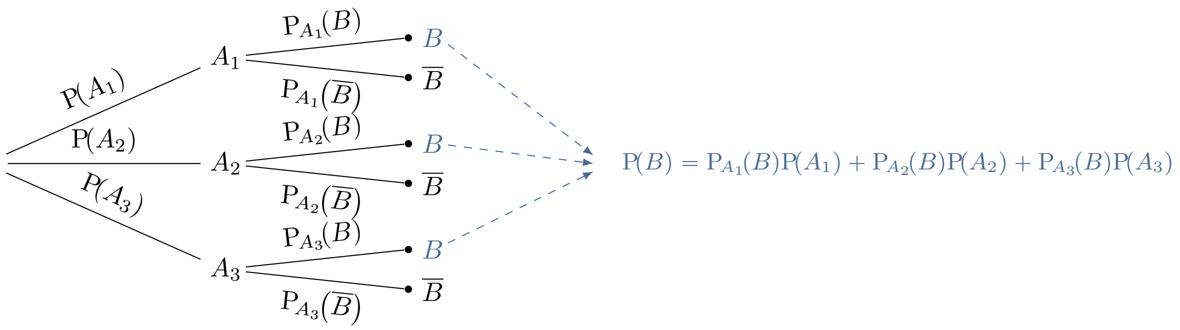
Il suffit alors d'appliquer la propriété d'additivité d'une probabilité à cette union d'événements deux à deux incompatibles. ■

Exemple 32 On reprend l'exemple 29. Quelle est la probabilité d'obtenir une boule noire au 2$^{\text{e}}$ tirage ?

Remarque 33 L'idée derrière la formule des probabilités totales est celle de l'étude de cas. Dans l'exemple précédent, on teste l'événement $\overline{B_2}$ selon les deux issues possibles du premier tirage : soit on a obtenu une boule noire (événement $\overline{B_1}$), soit on a obtenu une boule blanche (événement B_1).

- L'étude de cas se doit d'être exhaustive, *i.e.* inclure tous les cas possibles.
  $\hookrightarrow$ Ceci est assuré par le fait que la réunion des événements d'un système complet d'événements est Ω.
- Chaque cas n'empiète sur aucun autre.
  $\hookrightarrow$ Ceci est assuré par le fait que les événements d'un système complet d'événements sont 2 à 2 incompatibles.

Vous avez déjà rencontré la formule des probabilités totales auparavant, mais sans doute dans le contexte des *arbres de probabilité*. Avec la représentation suivante pour le cas $n = 3$:



Dès à présent, vous conservez naturellement le droit de penser en termes d'arbres de probabilités si cela vous aide, cependant un tel arbre ne sera plus considéré comme une preuve correctement formalisée. Autrement dit, votre rédaction ne doit pas s'appuyer sur les arbres, mais sur la formule des probabilités totales.

3.2.3 Formule de Bayes ou probabilité des causes

La formule qui suit relie les deux probabilités symétriques $\mathrm{P}_B(A)$ et $\mathrm{P}_A(B)$.

Théorème 34 – Formule de Bayes

(i) Si $A, B \in \mathscr{P}(\Omega)$ sont deux événements de probabilités NON NULLES, alors

$$\mathrm{P}_B(A) = \frac{\mathrm{P}(A)}{\mathrm{P}(B)}\mathrm{P}_A(B).$$

(ii) Si $(A_1, \dots, A_n)$ est un système complet d'événements et $B \in \mathscr{P}(\Omega)$ un événement de probabilité NON NULLE, alors

$$\forall j \in [\![1\,,n]\!], \quad \mathrm{P}_B(A_j) = \frac{\mathrm{P}(A_j)\mathrm{P}_{A_j}(B)}{\displaystyle\sum_{i=1}^{n} \mathrm{P}(A_i)\mathrm{P}_{A_i}(B)}.$$

Démonstration. **(i)** On a tout simplement par définition d'une probabilité conditionnelle $\mathrm{P}_B(A) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \dfrac{\mathrm{P}(A)\mathrm{P}_A(B)}{\mathrm{P}(B)}$.

(ii) est une conséquence de **(i)** et de l'écriture de $\mathrm{P}(B)$ via la formule des probabilités totales. ■

Exemple 35 On reprend l'exemple précédent. Le second tirage ayant donné une boule blanche, quelle est la probabilité que la première boule tirée ait été blanche ?

Remarque 36 – Formule des causes La formule de Bayes est également connue sous le nom de *formule des causes*. Si l'on reprend l'exemple précédent, l'événement B_2 est postérieur à l'événement B_1. On cherche à connaître la probabilité de B_1, qui est une cause possible de l'événement B_2, sachant que la conséquence B_2 s'est réalisée. Autrement dit, on cherche à mesurer l'influence que B_1 a eue dans la réalisation de B_2 et ainsi savoir si B_1 est une cause probable de B_2.

Exercice 37 On considère une population atteinte par une maladie rare touchant une personne sur 10 000. Un test de dépistage est proposé et donne les résultats suivants :

- si une personne est malade, le test est positif à 99% ;
- si une personne est saine, le test peut se révéler positif à hauteur de 0,1% (on parle de faux positif).

A-t-on intérêt à se fier aux résultats de ce test ?

4 Loi d'une variable aléatoire

Dans l'ensemble de cette section, (Ω, P) désigne un espace probabilisé fini.

4.1 Définitions et premières propriétés

Définition-théorème 38 – Loi d'une variable aléatoire

Soit X une variable aléatoire sur Ω. On appelle *loi de X (pour la probabilité* P*)* l'application

$$\mathrm{P}_X : \mathscr{P}(X[\Omega]) \longrightarrow [0\,,1], A \longmapsto \mathrm{P}(X \in A)$$

qui définit une probabilité sur $X[\Omega]$. Cette application est entièrement déterminée par la distribution de probabilités $(\mathrm{P}(X = x))_{x \in X[\Omega]}$ sur $X[\Omega]$, appelée la *distribution de probabilités de X*, en l'occurence

$$\forall A \in \mathscr{P}(X[\Omega]), \quad \mathrm{P}_X(A) = \mathrm{P}(X \in A) = \sum_{x \in A} \mathrm{P}(X = x).$$

Démonstration. Vérifions que P_X définit une probabilité sur $X[\Omega]$:

- pour tout $A \in \mathscr{P}(X[\Omega])$, $\mathrm{P}_X(A) = \mathrm{P}(X \in A) \in [0\,,1]$;
- $\mathrm{P}_X(X[\Omega]) = \mathrm{P}(X \in X[\Omega]) = \mathrm{P}(\Omega) = 1$;
- pour tout $A, B \in \mathscr{P}(X[\Omega])$ incompatibles,

$$\mathrm{P}_X(A \sqcup B) = \mathrm{P}(X \in A \sqcup B) = \mathrm{P}((X \in A) \sqcup (X \in B)) = \mathrm{P}(X \in A) + \mathrm{P}(X \in B) = \mathrm{P}_X(A) + \mathrm{P}_X(B).$$ ■

On définit ainsi une probabilité P_X directement sur l'ensemble $X[\Omega]$ (souvent plus simple que Ω) naturellement issue de la probabilité P déjà connue sur Ω. La loi de X mesure la vraisemblance des valeurs de X. Les valeurs possibles de X, tout en étant possibles, n'ont en effet pas forcément les mêmes chances de réalisation et c'est précisément cette information que P_X mesure.

✎ **En pratique** ✎ Expliciter la loi P_X d'une variable aléatoire réelle finie revient à déterminer :

(i) l'image $X[\Omega]$; **(ii)** la valeur de $\mathrm{P}(X = x)$, pour tout $x \in X[\Omega]$.

Exemple 39 Les lois de probabilité des variables aléatoires X_1 et S de l'exemple 9 sont résumées par les tableaux ci-dessous.

i	1	2	3	4	5	6
$\mathrm{P}(X_1 = i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

k	2	3	4	5	6	7	8	9	10	11	12
$\mathrm{P}(S = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Définition 40 – Loi conditionnelle d'une variable aléatoire

Soit X une variable aléatoire sur Ω et $A \in \mathscr{P}(\Omega)$ tel que $\mathrm{P}(A) \neq 0$. On appelle *loi conditionnelle de X sachant A (pour la probabilité* P*)* la loi de X pour la probabilité P_A, *i.e.* l'application

$$\mathscr{P}(X[\Omega]) \longrightarrow [0\,,1], B \longmapsto \mathrm{P}_A(X \in B).$$

À nouveau, la loi conditionnelle de X sachant A est entièrement déterminée par la distribution de probabilités $(\mathrm{P}_A(X = x))_{x \in X[\Omega]}$.

Exemple 41 La loi conditionnelle de la variable aléatoire X_1 de l'exemple 9 sachant $(S = 4)$ est donnée par le tableau ci-contre.

k	1	2	3	4	5	6
$\mathrm{P}_{(S=4)}(X_1 = k)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0	0

Le théorème suivant indique qu'il est possible de définir la loi d'une variable aléatoire *ex nihilo*, *i.e.* sans devoir décrire complètement l'expérience aléatoire sous-jacente. Ainsi, on peut escamoter l'espace probabilisé (Ω, P) au profit de la seule loi de la variable aléatoire. Ce résultat justifie la consistance de tous les énoncés de théorèmes et d'exercices qui, pour définir une variable aléatoire, nous en donnent seulement l'image et les probabilités des événements du système complet d'événements associé. Il sera ainsi largement mis à profit lors de la définition des lois usuelles au paragraphe suivant.

Théorème 42 – Définition implicite d'un espace probabilisé via une distribution de probabilités

Si I un ensemble fini et $(p_i)_{i \in I}$ une distribution de probabilités sur I, alors il existe un espace probabilisé fini (Ω, P) et une variable aléatoire X sur Ω d'image I pour lesquels $\mathrm{P}(X = i) = p_i$, pour tout $i \in I$.

Démonstration. Considérons $\Omega = I$, $X = \mathrm{Id}_I$ et définissons une probabilité sur Ω en posant, pour tout $i \in I$, $\mathrm{P}(\{i\}) = p_i$ (théorème 20). Comme voulu, on a, pour tout $i \in I$, $\mathrm{P}(X = i) = \mathrm{P}(\mathrm{Id}_I = i) = \mathrm{P}(\{i\}) = p_i$. ∎

Une variable aléatoire réelle étant avant tout une fonction à valeurs réelles, il est loisible d'opérer sur l'ensemble des variables aléatoires réelles. Ainsi, si X et Y sont des variables aléatoires réelles, il en va alors de même de $X + Y$, XY, λX, avec $\lambda \in \mathbb{R}$, etc. En toute généralité, intéressons nous à la loi image d'une variable aléatoire par une fonction f.

Définition-théorème 43 – Loi image d'une variable aléatoire par une fonction

Soit X une variable aléatoire réelle sur Ω, F un ensemble et $f : X[\Omega] \longrightarrow F$ une fonction. La loi $\mathrm{P}_{f(X)}$ de $f(X) = f \circ X$ est entièrement déterminée par f et la loi de X. Précisément,

$$\forall A \in \mathscr{P}(f[X[\Omega]]), \quad \mathrm{P}_{f(X)}(A) = \mathrm{P}(f(X) \in A) = \sum_{\substack{x \in X[\Omega] \\ f(x) \in A}} \mathrm{P}(X = x).$$

Démonstration. Pour tout $A \in \mathscr{P}(f[X[\Omega]])$, il suffit de remarquer que $(f(X) \in A)$ se réécrit :

$$(f(X) \in A) \cap \Omega = (f(X) \in A) \cap \left(\bigsqcup_{x \in X[\Omega]} (X = x) \right) = \bigsqcup_{x \in X[\Omega]} (f(X) \in A \text{ et } X = x) = \bigsqcup_{x \in X[\Omega]} (f(x) \in A \text{ et } X = x) = \bigsqcup_{\substack{x \in X[\Omega] \\ f(x) \in A}} (X = x).$$ ∎

Ce théorème montre en particulier que si deux variables aléatoires X et Y ont la même loi, les variables $f(X)$ et $f(Y)$ ont aussi la même loi pour toute fonction f pour laquelle la composition a un sens.

✎ **En pratique** ✎ **Détermination de la loi de $Y = f(X)$**

(i) On commence par déterminer l'image de Y, qui s'obtient comme l'image direct de l'image $X[\Omega]$ de X par la fonction f : $Y[\Omega] = f[X[\Omega]]$.

(ii) Pour chaque valeur $y \in Y[\Omega]$, on détermine $\mathrm{P}(Y = y)$, en décomposant l'événement $(Y = y)$ à l'aide du système complet d'événements $((X = x))_{x \in X[\Omega]}$ associé à X. En pratique, cela revient à déterminer les antécédents de y dans $X[\Omega]$ par la fonction f.

Exemple 44 Soit X une variable aléatoire à valeurs dans $[\![-2\,,2]\!]$ telle que $\mathrm{P}(X = x) = \dfrac{1}{5}$, pour tout $x \in [\![-2\,,2]\!]$ (X suit la loi uniforme sur $[\![-2\,,2]\!]$, cf. définition 45). Déterminer la loi de $Y = X^2 + 1$.

4.2 Lois usuelles

4.2.1 Loi uniforme

Définition 45 – Loi uniforme

Soit E un ensemble fini non vide et X une variable aléatoire à valeurs dans E.
On dit que X *suit la loi uniforme sur* E ou que X *est une variable uniforme sur* E, et on note $X \sim \mathscr{U}(E)$, lorsque P_X est la probabilité uniforme sur E, *i.e.* $\mathrm{P}(X = x) = \dfrac{1}{|E|}$, pour tout $x \in E$.

On définit ainsi correctement une loi de probabilité (cf. théorème 42), puisque

- pour tout $x \in E$, $\mathrm{P}(X = x) = \dfrac{1}{|E|} \geqslant 0$;
- $\displaystyle\sum_{x \in E} \mathrm{P}(X = x) = |E| \times \frac{1}{|E|} = 1$.

La loi uniforme est la loi pour laquelle les événements $(X = x)$, x décrivant E, sont équiprobables de probabilité $\dfrac{1}{|E|}$. Elle permet évidemment de modéliser les situations d'équiprobabilité.

Exemple 46

- Lorsque l'on choisit un entier X au hasard entre 1 et n, X suit la loi uniforme sur $[\![1\,,n]\!]$.
- Lorsque l'on lance deux fois successivement un dé équilibré à 6 faces, le couple (X_1, X_2) des valeurs obtenues suit la loi uniforme sur $[\![1\,,6]\!]^2$.
- La loi uniforme sur $\{-1, 1\}$ est appelée *loi de Rademacher*. Ainsi une variable aléatoire X suit la loi de Rademacher lorsqu'elle vérifie $\mathrm{P}(X = -1) = \mathrm{P}(X = 1) = \dfrac{1}{2}$.

Exemple 47 On choisit un entier X au hasard entre 1 et $2n$. Quelle est la loi conditionnelle de X sachant $(X \leqslant n)$? Quelle est la loi de $(-1)^X$?

4.2.2 Loi de Bernoulli

La loi de Bernoulli[†] apparait dans toute expérience de type « succès/échec ». Précisément, on appelle *épreuve de Bernoulli de paramètre* $p \in [0\,,1]$ toute expérience aléatoire à deux issues : une nommée « succès », de probabilité p, et l'autre « échec », de probabilité $1 - p$, avec $p \in [0\,,1]$.

Exemple 48 L'expérience aléatoire consistant à lancer une pièce équilibrée avec pour succès l'obtention de la face « pile » est une épreuve de Bernoulli de paramètre $p = 1/2$.

Définition 49 – Loi de Bernoulli

Soit $p \in [0\,,1]$. Une variable aléatoire X suit la *loi de Bernoulli de paramètre* p, ce que l'on note $X \sim \mathscr{B}(p)$, lorsque

$$X[\Omega] = \{0, 1\}, \qquad \mathrm{P}(X = 1) = p \quad \text{et} \quad \mathrm{P}(X = 0) = 1 - p.$$

Exemple fondamental. Pour tout espace probabilisé (Ω, P) et pour tout $A \in \mathscr{P}(\Omega)$, $\mathbb{1}_A \sim \mathscr{B}(\mathrm{P}(A))$.

On définit ainsi correctement une loi de probabilité, puisque $p \geqslant 0$, $1 - p \geqslant 0$ et $p + (1 - p) = 1$ (cf. théorème 42).

Une variable aléatoire X suit la loi de Bernoulli de paramètre p lorsqu'elle est liée à une épreuve de Bernoulli de paramètre p en associant la valeur 1 au « succès » et la valeur 0 à l'« échec ».

Exemple 50 Lors du lancer d'un dé, Alice gagne 1 euro lorsque le résultat est supérieur à 5 et rien sinon. On note G son gain. Alors la variable aléatoire G suit une loi de Bernoulli de paramètre $\frac{1}{3}$.
En effet, $G[\Omega] = \{0, 1\}$ et $\mathrm{P}(G = 1) = \mathrm{P}(\{5, 6\}) = \frac{2}{6} = \frac{1}{3}$. Ici le « succès » est l'événement « le résultat est supérieur à 5 ».

†. En l'honneur de Jacques Bernoulli (1654 à Bâle - 1705 à Bâle) mathématicien et physicien suisse ayant notamment contribué à l'étude des probabilités, des séries et de certaines équations différentielles.

Exemple 51 Soit $p \in [0\,,1]$ et $k \in \mathbb{N}^*$. Si $X \sim \mathscr{B}(p)$, alors $X^k \sim \mathscr{B}(p)$.
En effet, $X^k[\Omega] = \{0,1\}$ et $\mathrm{P}(X^k = 1) = \mathrm{P}(X = 1) = p$.

Les indicatrices jouent un rôle naturel en théorie des probabilités comme l'indique l'exemple suivant.

Exemple 52 On lance n fois un dé à 6 faces. Si l'on note A_k l'événement « On obtient 6 au k^{e} lancer », pour tout $k \in [\![1\,,n]\!]$, et N le nombre de 6 obtenus, alors $N = \mathbb{1}_{A_1} + \ldots + \mathbb{1}_{A_n}$.

Toute variable aléatoire qui représente un cardinal peut être exprimée comme une somme d'indicatrices, *i.e.* comme une somme de variables aléatoires suivant une loi de Bernoulli.

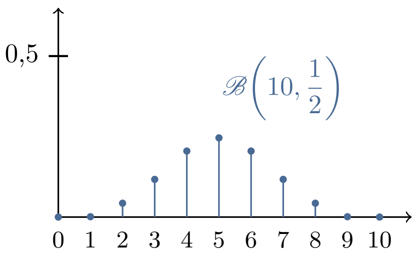
4.2.3 Loi binomiale

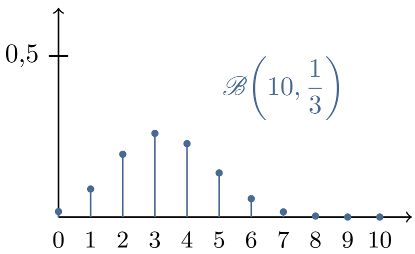
On se contente dans ce paragraphe de donner la définition *ex nihilo* d'une loi binomiale. Son interprétation probabiliste sera donnée à la section 5.1.2.

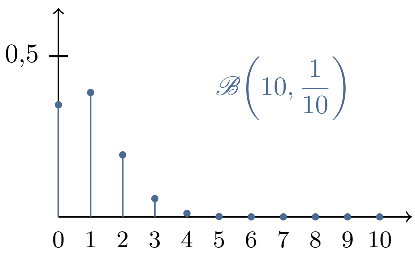
Définition 53 – Loi binomiale ou loi des tirages avec remise

Soit $n \in \mathbb{N}^*$ et $p \in [0\,,1]$. Une variable aléatoire X suit la *loi binomiale de paramètres n et p*, ce que l'on note $X \sim \mathscr{B}(n,p)$, lorsque

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







On définit bien ainsi une loi de probabilité (cf. théorème 42), puisque $\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \geqslant 0$, pour tout $k \in [\![0\,,n]\!]$, et, d'après la formule du binôme,

$$\sum_{k=0}^{n} \mathrm{P}(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (\cancel{p} + (1-\cancel{p}))^n = 1^n = 1.$$

Exemple 54 Pour toute variable aléatoire X suivant une loi binomiale de paramètre (n,p), la variable aléatoire $n - X$ suit une loi binomiale de paramètre $(n, 1-p)$.

5 Indépendance

Dans cette section, (Ω, P) désigne un espace probabilisé fini.

5.1 Événements indépendants

5.1.1 Indépendance d'une paire d'événements

Définition 55 – Paire d'événements indépendants

Deux événements A et B de $\mathscr{P}(\Omega)$ sont dits *indépendants (pour la probabilité* P) lorsque $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$.

Par définition d'une probabilité conditionnelle, on en déduit immédiatement la caractérisation suivante.

Théorème 56

Soit A et B deux événements de $\mathscr{P}(\Omega)$. Si $\mathrm{P}(A) \neq 0$, alors

$$A \text{ et } B \text{ sont indépendants (pour la probabilité P) si et seulement si } \mathrm{P}(B) = \mathrm{P}_A(B).$$

Remarque 57 La caractérisation précédente est limpide, puisqu'elle exprime que la probabilité de réalisation de l'événement B est la même que celle de l'événement « B sachant A », ce qui suggère bien l'absence d'influence de A sur B.

✎ En pratique ✎

- L'indépendance de deux événements peut se justifier par les modalités de l'expérience. Par exemple, lors de deux lancers successifs d'un dé, le résultat du premier jet n'influe pas celui du second, ainsi tous les événements liés uniquement au premier jet seront indépendants de ceux liés au second. On peut alors calculer aisément la probabilité d'une intersection de deux tels événements.
- Pour établir par le calcul l'indépendance de deux événements A et B, on calcule explicitement $\mathrm{P}(A)$ et $\mathrm{P}_B(A)$ (ou $\mathrm{P}(B)$ et $\mathrm{P}_A(B)$) et on vérifie que ces deux probabilités sont égales.

Exemple 58 On lance un dé cubique, ainsi $\Omega = \{1, 2, 3, 4, 5, 6\}$, et on considère les événements :

- A : « le résultat obtenu est inférieur à 2 » ;
- B : « le résultat obtenu est supérieur à 4 ».

On cherche à déterminer si A et B sont indépendants relativement à deux probabilités distinctes.

- **Cas 1 : dé équilibré.** L'univers Ω est donc muni de la probabilité uniforme P_1 déterminée par
$$\mathrm{P}_1(\{1\}) = \ldots = \mathrm{P}_1(\{6\}) = \frac{1}{6}.$$
Ainsi $\mathrm{P}_1(A) = \frac{2}{6} = \frac{1}{3}$ et $\mathrm{P}_1(B) = \frac{3}{6} = \frac{1}{2}$. Or A et B sont incompatibles, ainsi $\mathrm{P}_1(A \cap B) = 0 \neq \mathrm{P}_1(A)\mathrm{P}_1(B)$. Les événements A et B ne sont pas indépendants pour la probabilité P_1.
- **Cas 2 : dé pipé.** Le dé considéré permet d'obtenir 1 avec la probabilité 1, ainsi Ω est muni de la probabilité P_2 déterminée par
$$\mathrm{P}_2(\{1\}) = 1 \quad \text{et} \quad \mathrm{P}_2(\{2\}) = \ldots = \mathrm{P}_2(\{6\}) = 0.$$
On a donc $\mathrm{P}_2(A) = 1$ et $\mathrm{P}_2(B) = 0$. Or A et B sont incompatibles, ainsi $\mathrm{P}_2(A \cap B) = \mathrm{P}(\varnothing) = 0 = \mathrm{P}_2(A)\mathrm{P}_2(B)$. Les événements A et B sont donc indépendants pour la probabilité P_2.

La notion d'indépendance n'est pas intrinsèque aux événements, mais dépend fortement de la probabilité qui modélise l'expérience aléatoire considérée.

✖ Attention ! ✖ Il ne faut donc pas confondre, pour deux événements A et B, les deux notions suivantes :

- A et B sont incompatibles : notion intrinsèque aux événements et qui ne dépend d'aucune probabilité.
- A et B sont indépendants : notion qui dépend de la probabilité P choisie sur Ω.

Précisément,
$$A \text{ et } B \text{ incompatibles} \quad \not\Rightarrow \quad A \text{ et } B \text{ indépendants (pour P)},$$
comme l'illustre l'exemple précédent. Et
$$A \text{ et } B \text{ incompatibles} \quad \not\Leftarrow \quad A \text{ et } B \text{ indépendants (pour P)},$$
comme l'illustre l'exemple qui suit.

Exemple 59 On lance deux fois un dé à 6 faces, l'univers est $\Omega = [\![1\,,6]\!]^2$, et on considère les événements

- A : « le premier chiffre est pair » ;
- B : « le second chiffre est impair ».

Le dé est supposé équilibré et on munit Ω de la probabilité uniforme, notée P.
On a alors $\mathrm{P}(A) = \dfrac{|A|}{|\Omega|} = \dfrac{3 \times 6}{6 \times 6} = \dfrac{1}{2}$ et, de même, $\mathrm{P}(B) = \dfrac{1}{2}$. Or

$$\mathrm{P}(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{3 \times 3}{36} = \frac{1}{4} = \mathrm{P}(A)\mathrm{P}(B).$$

En effet, $A \cap B = \{2,4,6\} \times \{1,3,5\}$. Ainsi A et B sont indépendants (pour P) et ne sont pas incompatibles. Pour l'indépendance, on aurait plus simplement pu arguer du fait que les deux lancers sont indépendants.

Théorème 60 – Indépendance *vs* complémentaire

Si A et B sont deux événements indépendants, alors

(i) A et $\overline{B}$ sont indépendants ; **(ii)** $\overline{A}$ et B sont indépendants ; **(iii)** $\overline{A}$ et $\overline{B}$ sont indépendants.

Démonstration. D'après le point **(ii)** du théorème 15,

$$\mathrm{P}(A \cap \overline{B}) = \mathrm{P}(A \backslash B) = \mathrm{P}(A) - \mathrm{P}(A \cap B) \underset{\text{indép.}}{=} \mathrm{P}(A) - \mathrm{P}(A)\mathrm{P}(B) = \mathrm{P}(A)(1 - \mathrm{P}(B)) = \mathrm{P}(A)\mathrm{P}(\overline{B}),$$

ce qui établit **(i)**. La démonstration de **(ii)** est totalement similaire et **(iii)** résulte de **(i)** et **(ii)**. ∎

5.1.2 Indépendance mutuelle

Définition 61 – Indépendance mutuelle, indépendance deux à deux

Soit $A_1, \ldots, A_n \in \mathscr{P}(\Omega)$.

(i) Les événements $A_1, \ldots, A_n$ sont dits *deux à deux indépendants (pour la probabilité* P*)* lorsque

$$\forall (i,j) \in [\![1\,,n]\!]^2, \quad i \neq j \implies \mathrm{P}(A_i \cap A_j) = \mathrm{P}(A_i)\mathrm{P}(A_j).$$

(ii) Les événements $A_1, \ldots, A_n$ sont dits *(mutuellement) indépendants (pour la probabilité* P*)* lorsque

$$\forall\, I \subset [\![1\,,n]\!], \quad \mathrm{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathrm{P}(A_i).$$

Afin de saisir la nuance entre ces deux notions, détaillons la définition précédente pour une famille de 3 événements (A_1, A_2, A_3).

- Les événements A_1, A_2, A_3 sont deux à deux indépendants lorsque :

$$\mathrm{P}(A_1 \cap A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2), \quad \mathrm{P}(A_1 \cap A_3) = \mathrm{P}(A_1)\mathrm{P}(A_3) \quad \text{et} \quad \mathrm{P}(A_2 \cap A_3) = \mathrm{P}(A_2)\mathrm{P}(A_3).$$

- Les événements A_1, A_2, A_3 sont mutuellement indépendants lorsque :

$$\mathrm{P}(A_1 \cap A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2), \quad \mathrm{P}(A_1 \cap A_3) = \mathrm{P}(A_1)\mathrm{P}(A_3), \quad \mathrm{P}(A_2 \cap A_3) = \mathrm{P}(A_2)\mathrm{P}(A_3)$$
$$\text{ET} \quad \mathrm{P}(A_1 \cap A_2 \cap A_3) = \mathrm{P}(A_1)\mathrm{P}(A_2)\mathrm{P}(A_3).$$

Remarque 62 De par la définition même de la mutuelle indépendance, toute sous-famille d'une famille d'événements indépendants est une famille d'événements indépendants.

✖ **Attention !** ✖ $\boxed{\text{(Mutuellement) indépendants} \implies \text{Deux à deux indépendants}}$ – mais la réciproque est fausse !

Exemple 63 On considère à nouveau l'expérience aléatoire de l'exemple 59 consistant à lancer deux fois un dé et on considère les événements

• A : « le premier chiffre est pair » ; • B : « le second chiffre est impair » ; • C : « la somme des chiffres est paire ».

Les événements A, B et C sont deux à deux indépendants mais ne sont pas mutuellement indépendants.
En effet,

1. Démontrons que A, B et C sont deux à deux indépendants.
   - On a déjà établi l'indépendance de A et B.
   - $\mathrm{P}(A \cap C) = \dfrac{|A \cap C|}{|\Omega|} = \dfrac{3^2}{6^2} = \dfrac{1}{4}$, en effet $A \cap C = \{2, 4, 6\}^2$, et, de même, $\mathrm{P}(B \cap C) = \dfrac{1}{4}$.
   - La famille $(A, \overline{A})$ forme un système complet d'événements et, via la formule des probabilités totales,
   $$\mathrm{P}(C) = \mathrm{P}(A \cap C) + \mathrm{P}(\overline{A} \cap C) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$
   - On vérifie alors que $\mathrm{P}(A \cap C) = \mathrm{P}(A)\mathrm{P}(C)$ et $\mathrm{P}(B \cap C) = \mathrm{P}(B)\mathrm{P}(C)$.
2. Démontrons que A, B et C ne sont pas mutuellement indépendants.
   - $A \cap B \cap C = \varnothing$, ainsi $\mathrm{P}(A \cap B \cap C) = 0$.
   - Or $\mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C) = \dfrac{1}{2} \times \dfrac{1}{2} \times \dfrac{1}{2} = \dfrac{1}{8}$.

Le résultat suivant généralise celui du théorème 60.

Théorème 64 – Indépendance *vs* complémentaires

Soit $A_1, \dots, A_n \in \mathscr{P}(\Omega)$. Pour tout $i \in [\![1, n]\!]$, B_i désigne A_i ou $\overline{A_i}$.

(i) Si les événements $A_1, \dots, A_n$ sont deux à deux indépendants, alors les événements $B_1, \dots, B_n$ le sont également.

(ii) Si les événements $A_1, \dots, A_n$ sont (mutuellement) indépendants, alors les événements $B_1, \dots, B_n$ le sont également.

Démonstration.

(i) Résulte directement du théorème 60.

(ii) Il suffit de démontrer que les événements obtenus en niant un seul des A_i sont mutuellement indépendants. Démontrons alors sans perte de généralité que $A_1, \dots, A_{n-1}, \overline{A_n}$ sont mutuellement indépendants.
Soit I un sous-ensemble de $[\![1, n]\!]$ de cardinal supérieur à 2. Si $n \notin I$, il n'y a rien à démontrer, par hypothèses sur les $(A_i)_{1 \leqslant i \leqslant n-1}$. Si $n \in I$, posons $I = \{i_1, \dots, i_k, n\}$, où $k = \mathrm{Card}(I) - 1$. Alors, d'après la formule des probabilités totales, avec le système complet d'événements $(A_n, \overline{A_n})$,
$$\mathrm{P}\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \mathrm{P}\left(\left(\bigcap_{j=1}^{k} A_{i_j}\right) \cap A_n\right) + \mathrm{P}\left(\left(\bigcap_{j=1}^{k} A_{i_j}\right) \cap \overline{A_n}\right)$$
et, par indépendance mutuelle des $(A_i)_{1 \leqslant i \leqslant n}$,
$$\mathrm{P}\left(\left(\bigcap_{j=1}^{k} A_{i_j}\right) \cap \overline{A_n}\right) = \prod_{j=1}^{k} \mathrm{P}(A_{i_j}) - \left(\prod_{j=1}^{k} \mathrm{P}(A_{i_j})\right) \times \mathrm{P}(A_n) = \prod_{j=1}^{k} \mathrm{P}(A_{i_j}) \times (1 - \mathrm{P}(A_n)) = \left(\prod_{j=1}^{k} \mathrm{P}(A_{i_j})\right) \times \mathrm{P}(\overline{A_n}).$$

■

Exemple 65 Un jeu de 32 cartes a été truqué de telle sorte qu'une carte autre que l'as de pique a été remplacée par un second as de pique. On répète alors n fois avec remise l'expérience consistant à tirer simultanément 4 cartes. À partir de quelle valeur de n la probabilité de déceler l'escroquerie est-elle supérieure ou égale à $0,9$?

Expérience modèle associée à une loi binomiale

Un *schéma de Bernoulli de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$* est une expérience aléatoire consistant à répéter n fois de manière INDÉPENDANTE une même épreuve de Bernoulli de paramètre p. On observe ainsi une succession de succès et d'échecs mutuellement indépendants et on s'intéresse à la variable aléatoire X comptant le nombre de succès observés suite aux n répétitions. Quelle est sa loi ?

Son support est $X[\Omega] = [\![0\,,n]\!]$ et, pour tout $k \in [\![0\,,n]\!]$, l'événement $(X = k)$ est réalisé si et seulement si l'on a obtenu k succès et $n-k$ échecs. Or il y a $\binom{n}{k}$ manières de distribuer ces k succès parmi les n répétitions et chacune de ces possibilités a, par indépendance, la même probabilité $p^k(1-p)^{n-k}$ de se réaliser, ainsi $\mathrm{P}(X = k) = \binom{n}{k} p^k(1-p)^{n-k}$.

Lorsque l'on répète n fois indépendamment une expérience aléatoire à deux issues de probabilité p pour l'issue favorable, le NOMBRE D'ISSUES FAVORABLES suit la loi binomiale $\mathscr{B}(n, p)$.

Remarque 66 L'expression « loi des tirages AVEC remise » par laquelle on décrit parfois la loi binomiale trouve son explication dans l'exemple suivant : lorsque l'on tire AVEC remise – donc indépendamment – n boules dans une urne contenant des boules blanches en proportion p et des boules noires en proportion $1-p$, la variable aléatoire comptant le nombre de boules blanches tirées suit la loi binomiale $\mathscr{B}(n, p)$.

Exemple 67 On lance 5 fois un dé équilibré à 6 faces dont 2 blanches et 4 noires. Avec quelle probabilité obtient-on exactement 3 fois une face noire ?

5.2 Variables aléatoires indépendantes

Nous adaptons dans ce paragraphe le concept d'indépendance aux variables aléatoires.

Définition-théorème 68 – Variables aléatoires indépendantes

Soit $X_1, \ldots, X_n$ des variables aléatoires sur Ω.

- **Définition.** Les variables aléatoires $X_1, \ldots, X_n$ sont dites *(mutuellement) indépendantes* lorsque, pour tout $(A_1, \ldots, A_n) \in \mathscr{P}(X_1[\Omega]) \times \cdots \times \mathscr{P}(X_n[\Omega])$, les événements $(X_1 \in A_1), \ldots, (X_n \in A_n)$ sont indépendants, *i.e.*

$$\forall (A_1, \ldots, A_n) \in \mathscr{P}(X_1[\Omega]) \times \cdots \times \mathscr{P}(X_n[\Omega]), \quad \forall I \subset [\![1\,,n]\!], \quad \mathrm{P}\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \prod_{i\in I} \mathrm{P}(X_i \in A_i).$$

- **Caractérisation.** Les variables aléatoires $X_1, \ldots, X_n$ sont *(mutuellement) indépendantes* si et seulement si

$$\forall (x_1, \ldots, x_n) \in X_1[\Omega] \times \cdots \times X_n[\Omega], \quad \mathrm{P}((X_1 = x_1) \cap \cdots \cap (X_n = x_n)) = \mathrm{P}(X_1 = x_1) \cdots \mathrm{P}(X_n = x_n).$$

Démonstration. Le sens direct est clair (on prend $A_i = \{x_i\}$). Supposons donc que

$$\forall (x_1, \ldots, x_n) \in X_1[\Omega] \times \cdots \times X_n[\Omega], \quad \mathrm{P}((X_1 = x_1) \cap \cdots \cap (X_n = x_n)) = \mathrm{P}(X_1 = x_1) \cdots \mathrm{P}(X_n = x_n)$$

et montrons que les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes.

- *Étape 1.* Soit $(B_1, \ldots, B_n) \in \mathscr{P}(X_1[\Omega]) \times \cdots \times \mathscr{P}(X_n[\Omega])$, En notant $\mathbb{B} = B_1 \times \cdots \times B_n$,

$$\prod_{i=1}^{n} \mathrm{P}(X_i \in B_i) = \prod_{i=1}^{n} \sum_{x_i \in B_i} \mathrm{P}(X_i = x_i) = \sum_{(x_i)_{1\leqslant i\leqslant n} \in \mathbb{B}} \prod_{i=1}^{n} \mathrm{P}(X_i = x_i) = \sum_{(x_i)_{1\leqslant i\leqslant n} \in \mathbb{B}} \mathrm{P}\left(\bigcap_{1\leqslant i\leqslant n} (X_i = x_i)\right)$$

$$= \mathrm{P}\left(\bigsqcup_{(x_i)_{1\leqslant i\leqslant n} \in \mathbb{B}} \ \bigcap_{1\leqslant i\leqslant n} (X_i = x_i)\right) = \mathrm{P}((X_1, \ldots, X_n) \in \mathbb{B}) = \mathrm{P}\left(\bigcap_{1\leqslant i\leqslant n} (X_i \in B_i)\right).$$

- *Étape 2.* Soit $(A_1, \ldots, A_n) \in \mathscr{P}(X_1[\Omega]) \times \cdots \times \mathscr{P}(X_n[\Omega])$ et I une partie de $[\![1\,,n]\!]$. Pour tout $i \in [\![1\,,n]\!]$, posons

$$B_i = \begin{cases} A_i & \text{si } i \in I \\ X_i[\Omega] & \text{sinon.} \end{cases}$$

D'après le point précédent,

$$\mathrm{P}\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \mathrm{P}\left(\bigcap_{1\leqslant i\leqslant n} (X_i \in B_i)\right) = \prod_{i=1}^{n} \mathrm{P}(X_i \in B_i) = \prod_{i\in I} \mathrm{P}(X_i \in A_i).$$

■

L'indépendance des variables aléatoires étant définie comme une indépendance d'événements, toute sous-famille d'une famille de variables aléatoires indépendantes est une famille de variables aléatoires indépendantes.

✖ **ATTENTION !** ✖ À nouveau, (Mutuellement) indépendants $\Longrightarrow$ Deux à deux indépendants – mais la réciproque est fausse !

Exemple 69 On lance un dé équilibré à 6 faces deux fois et on note X_1 (resp. X_2) la face obtenue au premier (resp. second) lancer. Avec quelle probabilité obtient-on les deux fois une face impaire ?

En effet, par hypothèse, les variables aléatoires X_1 et X_2 sont indépendantes et uniformes sur $[\![1\,,6]\!]$, ainsi

$$\mathrm{P}(X_1 \text{ et } X_2 \text{ impaires}) = \mathrm{P}(X_1 \text{ impaire})\mathrm{P}(X_2 \text{ impaire}) = \left(\frac{|\{1,3,5\}|}{|[\![1\,,6]\!]|}\right)^2 = \frac{1}{4}.$$

Exemple 70 Si $p_1,\ldots,p_n \in [0\,,1]$ et si $X_1,\ldots,X_n$ sont des variables de Bernoulli indépendantes de paramètres respectifs $p_1,\ldots,p_n$, alors $X_1\cdots X_n \sim \mathscr{B}(p_1\cdots p_n)$.

En effet, la variable aléatoire $X_1\cdots X_n$ est à valeurs dans $\{0,1\}$ et

$$\mathrm{P}(X_1\cdots X_n = 1) = \mathrm{P}((X_1 = 1)\cap\cdots\cap(X_n = 1)) \overset{\text{indép.}}{=} \mathrm{P}(X_1 = 1)\cdots\mathrm{P}(X_n = 1) = p_1\cdots p_n.$$

À l'instar du théorème 42, le théorème suivant garantit la consistance des énoncés affirmant l'existence d'une famille de variables aléatoires INDÉPENDANTES de lois données.

Théorème 71 – Existence d'une famille finie de variables aléatoires indépendantes de lois prescrites

Soit E et F deux ensembles finis. Si $(p_e)_{e\in E}$ et $(q_f)_{f\in F}$ sont deux distributions de probabilités sur E et F respectivement, alors il existe un espace probabilisé fini (Ω,P) et des variables aléatoires INDÉPENDANTES X et Y sur Ω d'images respectives E et F telles que

$$\forall e\in E,\quad \mathrm{P}(X = e) = p_e \qquad \text{et} \qquad \forall f\in F,\quad \mathrm{P}(Y = f) = q_f.$$

Cet énoncé se généralise au cas d'un nombre fini quelconque de variables aléatoires.

Démonstration. Étant donné que $\displaystyle\sum_{(e,f)\in E\times F} p_e q_f = \sum_{e\in E} p_e \times \sum_{f\in F} q_f = 1$, la famille $(p_e q_f)_{(e,f)\in E\times F}$ est une distribution de probabilités sur $E\times F$. Le théorème 42 garantit alors l'existence d'un espace probabilisé (Ω,P) et d'une variable aléatoire (X,Y) sur Ω d'image $E\times F$ telle que, pour tout $(e,f)\in E\times F$, $\mathrm{P}((X,Y) = (e,f)) = p_e q_f$. On alors

- pour tout $e\in E$, d'après la formule des probabilités totales avec le système complet d'événements $((Y = f))_{f\in F}$,

$$\mathrm{P}(X = e) = \sum_{f\in F}\mathrm{P}((X = e)\cap(Y = f)) = \sum_{f\in F} p_e q_f = p_e$$

  et de même, pour tout $f\in F$, $\mathrm{P}(Y = f) = q_f$.
- X et Y sont indépendants, puisque $\mathrm{P}((X = e)\cap(Y = f)) = p_e q_f = \mathrm{P}(X = e)\mathrm{P}(Y = f)$, pour tout $(e,f)\in E\times F$. ∎

Théorème 72 – Lemme des coalitions

Soit E et F deux ensembles, $X_1,\ldots,X_n$ des variables aléatoires sur Ω ainsi que $f : (X_1,\ldots,X_m)[\Omega] \longrightarrow E$ et $g : (X_{m+1},\ldots,X_n)[\Omega] \longrightarrow F$ deux fonctions, avec $m\in[\![1\,,n]\!]$. Si les variables aléatoires $X_1,\ldots,X_n$ sont indépendantes, alors les coalitions $f(X_1,\ldots,X_m)$ et $g(X_{m+1},\ldots,X_n)$ le sont aussi.

Cet énoncé se généralise au cas d'un nombre fini quelconque de coalitions.

Démonstration. Notons $X = (X_1,\ldots,X_m)$ et $Y = (X_{m+1},\ldots,X_n)$.

- *Étape 1.* Indépendance de X et Y.
  Pour tous $x = (x_1,\ldots,x_m)\in X_1[\Omega]\times\cdots\times X_m(\Omega)$ et $y = (x_{m+1},\ldots,x_n)\in X_{m+1}(\Omega)\times\cdots\times X_n[\Omega]$,

$$\begin{aligned}\mathrm{P}((X = x)\cap(Y = y)) &= \mathrm{P}\left(\bigcap_{1\leqslant i\leqslant n}(X_i = x_i)\right) \overset{\text{indép.}}{=} \prod_{1\leqslant i\leqslant n}\mathrm{P}(X_i = x_i)\\ &\overset{\text{indép.}}{=} \mathrm{P}\left(\bigcap_{1\leqslant i\leqslant m}(X_i = x_i)\right)\mathrm{P}\left(\bigcap_{m+1\leqslant i\leqslant n}(X_i = x_i)\right) = \mathrm{P}(X = x)\mathrm{P}(Y = y).\end{aligned}$$

- *Étape 2.* Indépendance de $f(X)$ et $g(Y)$. Pour tous $a\in f(X)[\Omega]$ et $b\in g(Y)[\Omega]$,

$$\begin{aligned}\mathrm{P}((f(X) = a)\cap(g(Y) = b)) &= \mathrm{P}\big((X\in f^{-1}[\{a\}])\cap(Y\in g^{-1}[\{b\}])\big) \overset{\text{indép.}}{=} \mathrm{P}\big(X\in f^{-1}[\{a\}]\big)\mathrm{P}\big(Y\in g^{-1}[\{b\}]\big)\\ &= \mathrm{P}(f(X) = a)\mathrm{P}(g(Y) = b).\end{aligned}$$

∎

Même si cet énoncé peut sembler compliqué au premier abord, il ne fait qu'exprimer par exemple que si trois variables aléatoires réelles X, Y et Z définies sur Ω sont indépendantes, alors il en va de même de $X+Y$ et Z^2.

6 Couple de variables aléatoires

Dans l'ensemble de cette section, (Ω, P) désigne un espace probabilisé fini.

6.1 Loi conjointe et lois marginales

Définition-théorème 73 – Couple de variables aléatoires, loi conjointe et lois marginales

Soit X et Y deux variables aléatoires sur Ω.

- L'application $\omega \longmapsto (X(\omega), Y(\omega))$ est une variable aléatoire sur Ω notée (X, Y) et souvent qualifiée de *vecteur aléatoire*.
- On appelle *loi conjointe (ou loi mutuelle) de* (X, Y), notée $\mathrm{P}_{(X,Y)}$, la loi de la variable aléatoire (X, Y), soit la donnée des réels
$$\mathrm{P}((X = x) \cap (Y = y)) \quad \text{pour tout } (x, y) \in X[\Omega] \times Y[\Omega].$$
$P_{(X,Y)}$ définit alors une loi de probabilité sur $X[\Omega] \times Y[\Omega]$.
- On appelle *première (resp. deuxième) loi marginale de* (X, Y) la loi P_X de X (resp. la loi P_Y de Y).
- L'ensemble $\{(X = x) \cap (Y = y)\}_{(x,y)\in X[\Omega]\times Y[\Omega]}$ forme un système complets d'événements, appelé le *système complets d'événements associé à* (X, Y).

Démonstration. Conséquence des théorèmes 11 et 38. ■

Remarque 74

- Pour le vecteur aléatoire (X, Y), on considère usuellement pour image le produit $X[\Omega] \times Y[\Omega]$ des images des variables aléatoires X et Y plutôt que l'image à proprement parler de (X, Y), à savoir
$$(X, Y)[\Omega] = \{(X(\omega), Y(\omega)) \mid \omega \in \Omega\} \subset X[\Omega] \times Y[\Omega].$$
- La probabilité $\mathrm{P}((X = x) \cap (Y = y))$ est aussi notée $\mathrm{P}(X = x \text{ et } Y = y)$ ou $\mathrm{P}(X = x, Y = y)$.

Exemple 75 Une urne contient trois boules numérotées 0 et une boule numérotée 1. On tire successivement et sans remise 2 boules. Le couple de numéros ainsi obtenu est noté (N_1, N_2). Quelle est la loi de (N_1, N_2) ?

D'après la formule des probabilités composées,
$$\mathrm{P}((N_1 = 0) \cap (N_2 = 0)) = \mathrm{P}(N_1 = 0)\mathrm{P}_{(N_1=0)}(N_2 = 0) = \frac{3}{4} \times \frac{2}{3} = \frac{1}{2}$$
et de la même manière $\quad \mathrm{P}((N_1 = 0) \cap (N_2 = 1)) = \frac{3}{4} \times \frac{1}{3} = \frac{1}{4}$,

$\mathrm{P}((N_1 = 1) \cap (N_2 = 0)) = \frac{1}{4} \times \frac{3}{3} = \frac{1}{4} \quad$ et $\quad \mathrm{P}((N_1 = 1) \cap (N_2 = 1)) = \frac{1}{4} \times \frac{0}{3} = 0.$

$N_1 \backslash N_2$	0	1
0	$\frac{1}{2}$	$\frac{1}{4}$
1	$\frac{1}{4}$	0

On vérifie évidemment que $\frac{1}{2} + \frac{1}{4} + \frac{1}{4} + 0 = 1$.

Le théorème suivant indique qu'il est toujours possible de déduire les lois marginales à partir de la loi conjointe.

Théorème 76 – Calcul des lois marginales à partir de la loi conjointe

Soit X et Y deux variables aléatoires sur Ω. La loi conjointe $P_{(X,Y)}$ du couple (X, Y) détermine entièrement ses lois marginales P_X et P_Y. Précisément,
$$\forall x \in X[\Omega], \qquad \mathrm{P}(X = x) = \sum_{y\in Y[\Omega]} \mathrm{P}((X = x) \cap (Y = y))$$
$$\text{et} \quad \forall y \in Y[\Omega], \qquad \mathrm{P}(Y = y) = \sum_{x\in X[\Omega]} \mathrm{P}((X = x) \cap (Y = y)).$$

Démonstration. Conséquence immédiate de la formule des probabilités totales, appliquée avec les systèmes complets d'événements $((X = x))_{x\in X[\Omega]}$ et $((Y = y))_{y\in Y[\Omega]}$. ■

✎ **En pratique** ✎ Lorsque X et Y sont finies, on représente parfois la loi de (X, Y) sous forme d'un tableau à deux entrées donnant $\mathrm{P}((X = x) \cap (Y = y))$ en fonction des valeurs possibles de $x \in X[\Omega]$ en lignes et de $y \in Y[\Omega]$ en colonnes. La figure ci-dessous illustre alors le calcul de P_X et P_Y à partir de $P(X, Y)$. On a noté $X[\Omega] = \{x_1, \ldots, x_m\}$, $Y[\Omega] = \{y_1, \ldots, y_n\}$ et $p_{ij} = \mathrm{P}((X = x) \cap (Y = y))$, pour tout $(i, j) \in [\![1\,, m]\!] \times [\![1\,, n]\!]$.

$X \backslash Y$	y_1	y_2	$\ldots$	y_n	
x_1	p_{11}	p_{12}	$\ldots$	p_{1n}	$\longrightarrow \sum_{j=1}^{n} p_{1j} = \mathrm{P}(X = x_1)$
x_2	p_{21}	p_{22}	$\ldots$	p_{2n}	
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_m	p_{m1}	p_{m2}	$\ldots$	p_{mn}	$\longrightarrow \sum_{j=1}^{n} p_{mj} = \mathrm{P}(X = x_m)$

Première loi marginale P_X de (X, Y).

$$\mathrm{P}(Y = y_1) = \sum_{i=1}^{n} p_{i1} \qquad \cdots \qquad \sum_{i=1}^{n} p_{in} = \mathrm{P}(Y = y_n)$$

Deuxième loi marginale P_Y de (X, Y).

Remarque 77 – Généralisation au cas de n variables aléatoires Les notions précédentes se généralisent au cas d'un n-uplet $(X_1, \ldots, X_n)$ de variables aléatoires. On définit leur loi conjointe par la donnée de

$$\mathrm{P}\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) \qquad \text{pour tout } (x_i)_{1 \leqslant i \leqslant n} \in \prod_{i=1}^{n} X_i[\Omega].$$

Les lois marginales sont toujours les lois des variables aléatoires X_i.

En revanche, les lois marginales ne sont généralement pas suffisantes pour déterminer la loi conjointe d'un couple de variables aléatoires. On a besoin d'une information plus précise concernant le lien entre ces variables.

Définition 78 – Lois conditionnées

Soit (X, Y) un vecteur aléatoire sur Ω.

- Pour tout $x \in X[\Omega]$ tel que $\mathrm{P}(X = x) \neq 0$, on appelle *loi de Y conditionnée à $(X = x)$* la donnée des réels

$$\mathrm{P}_{(X=x)}(Y = y) = \frac{\mathrm{P}((X = x) \cap (Y = y))}{\mathrm{P}(X = x)} \qquad \text{pour tout } y \in Y[\Omega].$$

- Pour tout $y \in Y[\Omega]$ tel que $\mathrm{P}(Y = y) \neq 0$, on appelle *loi de X conditionnée à $(Y = y)$* la donnée des réels

$$\mathrm{P}_{(Y=y)}(X = x) = \frac{\mathrm{P}((X = x) \cap (Y = y))}{\mathrm{P}(Y = y)} \qquad \text{pour tout } x \in X[\Omega].$$

✎ **En pratique** ✎ Ainsi, en vertu de la formule des probabilités composées, la loi conjointe d'un vecteur aléatoire (X, Y) peut s'obtenir par la connaissance :

- soit de la loi marginale de X et, pour tout $x \in X[\Omega]$, de la loi de Y conditionnée à $(X = x)$;
- soit de la loi marginale de Y et, pour tout $y \in Y[\Omega]$, de la loi de X conditionnée à $(Y = y)$.

Remarque 79

- En général, les lois marginales du couple (X, Y) ne déterminent pas sa loi conjointe, mais cela devient le cas lorsque X et Y sont indépendantes. Le cas échéant, on a en effet

$$\forall (x, y) \in X[\Omega] \times Y[\Omega], \quad \mathrm{P}((X = x) \cap (Y = y)) = \mathrm{P}(X = x)\mathrm{P}(Y = y).$$

En particulier, lorsque deux variables aléatoires sont indépendantes, les « lignes » (resp. « colonnes ») de leur loi conjointe sont toutes proportionnelles. Plus généralement, on retiendra que

Les variables $X_1, \ldots, X_n$ sont indépendantes si et seulement si la distribution de probabilités de $(X_1, \ldots, X_n)$ se calcule par simple produit à partir des distributions de probabilités de $X_1, \ldots, X_n$.

- Observons que, pour tout $x \in X[\Omega]$ tel que $\mathrm{P}(X = x) \neq 0$, on a l'équivalence

$$\mathrm{P}((X = x) \cap (Y = y)) = \mathrm{P}(X = x)\mathrm{P}(Y = y) \iff \mathrm{P}(Y = y) = \frac{\mathrm{P}((X = x) \cap (Y = y))}{\mathrm{P}(X = x)} = \mathrm{P}_{(X=x)}(Y = y).$$

  Ainsi la loi de Y sachant $(X = x)$ ne dépend aucunement de l'événement$(X = x)$, ce qui traduit que la connaissance de X n'apporte aucune information concernant Y, et inversement. C'est bien là la signification de l'indépendance.

Remarque 80 Il y a un moyen simple de repérer des variables aléatoires qui ne sont pas indépendantes. En effet, si X et Y sont indépendantes, pour tout $(x, y) \in X[\Omega] \times Y[\Omega]$, on dispose de l'implication

$$\mathrm{P}(X = x) \neq 0 \text{ et } \mathrm{P}(Y = y) \neq 0 \implies \mathrm{P}((X = x) \cap (Y = y)) = \mathrm{P}(X = x)\mathrm{P}(Y = y) \neq 0.$$

Ainsi, un vecteur aléatoire dont la loi conjointe contient des « termes » nuls a peu de chance d'être composé de variables aléatoires indépendantes.

6.2 Fonction d'un couple de variables aléatoires

Nous nous intéressons dans ce paragraphe à un problème courant, mais peu commode à résoudre en général. Considérons deux variables aléatoires X et Y, et $Z = f(X, Y)$, où f est une fonction sur $X[\Omega] \times Y[\Omega]$. Les cas les plus fréquents sont la somme $X + Y$, le produit XY ou encore le maximum $\max(X, Y)$ et le minimum $\min(X, Y)$.

Théorème 81 – Loi de $f(X, Y)$

Soit X et Y deux variables aléatoires sur Ω et $f : X[\Omega] \times Y[\Omega] \longrightarrow F$ une fonction, avec F un ensemble. La loi $\mathrm{P}_{f(X,Y)}$ de la variable aléatoire $f(X, Y)$ est entièrement déterminée par f et la loi conjointe de (X, Y). Précisément,

$$\forall z \in f(X, Y)[\Omega], \quad \mathrm{P}(f(X, Y) = z) = \sum_{\substack{(x,y)\in X[\Omega]\times Y[\Omega] \\ z=f(x,y)}} \mathrm{P}((X = x) \cap (Y = y)).$$

Ce résultat se généralise aux familles d'un nombre fini quelconque de variables aléatoires.

La formule pour la loi est toutefois tout sauf pratique ! Le cas classique d'utilisation est celui de la somme de deux variables aléatoires. En effet, si $Z = X + Y$, alors

$$\forall z \in Z[\Omega], \quad \mathrm{P}(Z = z) = \sum_{\substack{(x,y)\in X[\Omega]\times Y[\Omega] \\ z=x+y}} \mathrm{P}((X = x) \cap (Y = y)) = \sum_{x\in X[\Omega]} \mathrm{P}((X = x) \cap (Y = z - x)),$$

et la somme ne porte plus que sur un seul indice. Observons que $\mathrm{P}((X = x) \cap (Y = z - x))$ est nul dès que $z-x \notin Y[\Omega]$.

Exemple 82 Soit X et Y deux variables aléatoires indépendantes suivant toutes les deux la loi uniforme sur $[\![1\,,n]\!]$. Déterminons la loi de $Z = X + Y$, dite *loi triangulaire*, et la loi de l'écart $|X - Y|$.

Théorème 83 – Sommes de variables aléatoires indépendantes de lois binomiales

Soit $p \in [0\,,1]$, $m, n \in \mathbb{N}^*$ et $X, Y, X_1, \dots, X_n$ des variables aléatoires définies sur un même espace probabilisé fini.

(i) Si $X \sim \mathscr{B}(m, p)$ et $Y \sim \mathscr{B}(n, p)$, et si X et Y sont indépendantes, alors $X + Y \sim \mathscr{B}(m + n, p)$.

(ii) Si $X_i \sim \mathscr{B}(p)$, pour tout $i \in [\![1\,,n]\!]$, et si $X_1, \dots, X_n$ sont indépendantes, alors $\sum_{i=1}^{n} X_i \sim \mathscr{B}(n, p)$.

Démonstration. ... ■

Le résultat précédent est tout à fait naturel, il ne fait qu'exprimer que si X compte le nombre de succès lors de m répétitions d'une même épreuve de Bernoulli de probabilité de succès p et Y le nombre de succès obtenus après n nouvelles répétitions, alors la variable aléatoire $X + Y$ compte le nombre de succès lors de $m + n$ répétitions.

Compétences à acquérir

- Écrire un événement par opérations à partir d'autres événements : exercices 1 et 5.
- Exploiter les propriétés d'une probabilité : exercices 2 et 4.
- Modéliser une expérience aléatoire par un espace probabilisé fini adapté : exercices 5 à 11.
- Utiliser la formule des probabilités composées : exercices 12 et 17.
- Utiliser la formule des probabilités totales : exercices 14, 15, 16, 18, 32, 33 et 43.
- Utiliser la formule de Bayes : exercices 13 et 14.
- Déterminer la loi d'une variable aléatoire : exercices 17 à 20, 35, 36, 38 et 40.
- Reconnaitre les situations modéliser par des lois usuelles et les utiliser : exercices 21 à 23.
- Établir que des événements sont indépendants : exercices 26, 28 et 32.
- Utiliser l'indépendance d'événements : exercices 27 à 31, 33 et 34.
- Établir que des variables aléatoires sont indépendantes : exercice 39.
- Utiliser l'indépendance de variables aléatoires : exercices 35, 37, 38 et 40 à 43.
- Déterminer la loi conjointe/les lois marginales d'un couple de variables aléatoires : exercices 36 et 37.

Quelques résultats classiques :

- Produit de variables de Bernoulli indépendantes (exemple 70).
- Loi triangulaire (sommes de deux variables uniformes indépendantes) (exemple 82).
- Sous-additivité finie (exercice 2).
- Paradoxe des anniversaires (exercice 11).
- Urne de Pólya (exercice 18).
- Absence de piles consécutifs (exercice 33).
- Loi du minimum d'une famille de variables aléatoires indépendantes (exercice 40).