I.6 Variables indépendantes

Proposition I.24 (Variables aléatoires discrètes indépendantes)**.** *Soit X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E_1 et E_2 respectivement. Les propositions suivantes sont équivalentes:*
(i) Pour tout $(A, B) \in \mathcal{P}(E_1) \times \mathcal{P}(E_2)$ $\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B)$
(ii) Pour tout $(A, B) \in \mathcal{P}(E_1) \times \mathcal{P}(E_2)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, i.e.

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B);$$

(iii) Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, i.e.

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y);$$

Dans ce cas, on dit que X et Y sont indépendantes, et on note $X \perp Y$.

Remarque I.25 (15.1.28)**.** *Conformément à la remarque faite lors de la définition de la loi d'une variable aléatoire, on peut se contenter de considérer $(A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega))$.*

Proposition I.26 (Famille de variables mutuellement indépendantes)**.** *Soit $(X_k)_{k\in K}$ une famille quelconque de variables aléatoires discrètes sur $(\Omega, \mathcal{T}, \mathbb{P})$, X_k étant à valeurs dans E_k. Les propositions suivantes sont équivalentes:*
(i) pour toute famille $(A_k)_{k\in K}$ telle que pour tout $k \in K$, $A_k \subset E_k$, les événements $[X_k \in A_k]$, $k \in K$, sont indépendants;
(ii) pour tout $J \in \mathcal{P}_f(K)$ et toute famille $(A_j)_{j\in J}$ telle que pour tout $j \in J$, $A_j \subset E_j$,

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

(iii) pour toute famille $(x_k)_{k\in K}$ telle que pour tout $k \in K$, $x_k \in X_k(\Omega)$, les événements $\{X_k = x_k\}$, $k \in K$, sont indépendants ;
(iv) pour tout $J \in \mathcal{P}_f(K)$ et toute famille $(x_j)_{j\in J}$ telle que pour tout $j \in J$, $x_j \in X_j(\Omega)$,

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

Dans ce cas, on dit que les variables aléatoires X_k, $k \in K$, sont indépendantes (ou mutuellement indépendantes).

Proposition I.27 (Sous-famille d'une famille de variables indépendantes)**.** *Si la famille $(X_k)_{k\in K}$ est une famille de variables aléatoires discrètes indépendantes, alors, pour tout $J \subset K$, $(X_j)_{j\in J}$ est une famille de variables aléatoires discrètes indépendantes.*

Corollaire I.28 (Indépendance 2 à 2)**.** *Si $(X_k)_{k\in K}$ est une famille de variables aléatoires discrètes indépendantes, alors les X_k sont 2 à 2 indépendants. La réciproque est fausse.*

En revanche, on peut caractériser l'indépendance par l'indépendance des sous-familles finies.

Proposition I.29 (Caractérisation de l'indépendance par les familles finies)**.** *Soit $(X_k)_{k\in K}$ une famille de variables aléatoires discrètes. Les propositions suivantes sont équivalentes:*
(i) $(X_k)_{k\in K}$ est une famille de variables indépendantes ; (ii) pour tout $J \in \mathcal{P}_f(K)$, $(X_j)_{j\in J}$ est une famille de variables indépendantes.

Proposition I.30 (Fonctions de variables indépendantes)**.** *Soit $X \perp Y$ deux v.a.r.d. indépendantes, et f et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Alors $f(X) \perp g(Y)$.*

Ceci est un cas particulier du résultat plus général ci-dessous :

Théorème I.31 (Lemme des coalitions, pour 2 coalitions)**.** *Soit $1 \leq m < n$, et $(X_1, \dots, X_m, \dots X_n)$ une famille de variables aléatoires discrètes indépendantes. Soit f et g deux fonctions définies sur $X_1(\Omega) \times \dots \times X_m(\Omega)$ et $X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$ respectivement. Alors $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.*

Éléments de preuve. Démonstration un peu technique, en exprimant tout en fonction des répartitions de probabilités des X_i. □

Par récurrence, on en déduit la version générale, qu'on exprime volontairement sous une forme un peu lâche.

Théorème I.32 (Lemme des coalitions)**.** *Soit $(X_k)_{k \in K}$ une famille de variables aléatoires discrètes indépendantes, et $(K_i)_{i \in I}$ une famille de sous-ensembles finis deux à deux disjoints de K. Soit $(Y_i)_{i \in I}$ une famille de variables aléatoires, telle que pour tout $i \in I$, Y_i s'écrive comme fonction des X_k, $k \in K_i$ (et pas des autres). Alors $(Y_i)_{i \in I}$ est une famille de variables aléatoires discrètes indépendantes.*

Éléments de preuve. Le cas K infini et I fini s'obtient par récurrence à partir du lemme des 2 coalitions (seul un nombre fini de X_k intervenant, puisque les parties K_i sont finies).
Le cas I infini provient alors du fait que l'indépendance d'une famille équivaut à l'indépendance de toutes ses sous-familles finies. □

On admet le résultat suivant, qui permet d'assurer l'existence d'espaces probabilisés modélisant des situations un peu complexes, notamment infinies. Par exemple, la modélisation précise d'une succession infinie de tirages à Pile ou Face indépendants n'est pas chose facile. L'univers Ω est naturellement $\{0,1\}^{\mathbb{N}}$, mais cet ensemble est en bijection avec $\mathbb{R}$, et le choix de la tribu $\mathcal{P}(\Omega)$ ne permet pas ici de modéliser correctement le tirage à pile ou face.
En effet, on peut montrer que toute mesure de probabilité sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mathbb{P})$ est déterminée par une répartition de probabilités discrète. Or, pour tout $\omega \in \Omega$, on obtient alors dans cet espace probabilisé $\mathbb{P}(\omega) = 0$ (par le même calcul que la probabilité de n'obtenir que des Piles). Ceci contredit le fait que $\mathbb{P}$ soit déterminée par une répartition de probabilités discrète.
La construction d'un modèle convenable est difficile, et n'est pas un objectif du programme. Nous utiliserons donc le théorème suivant, admis, pour justifier l'existence d'un modèle convenable.

Théorème I.33 (Espace probabilisé modélisant une suite de lois, admis)**.** *Soit E un ensemble et $(P_n)_{n \in \mathbb{N}}$ une suite de lois de probabilités discrètes sur E (ce qui équivaut à la donnée d'une suite de distributions de probabilités discrètes).*
Alors il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et une suite $(X_n)_{n \in \mathbb{N}}$ de variables indépendantes, définies sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E, et telle que

$$\forall n \in \mathbb{N}, \mathbb{P}_{X_n} = P_n.$$

Éléments de preuve. Démonstration hors-programme. □

Il est fréquent, pour étudier les caractéristiques d'une expérience, de la répéter une infinité de fois dans les mêmes conditions, de façon indépendante, et d'observer le résultat sous forme de valeur prise par une variable aléatoire.

Définition I.34 (variables i.i.d.)**.** *Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires. On dit que la suite (X_n) est indépendante identiquement distribuée (en abrégé officiel i.i.d.), si les X_n sont indépendants, et suivent la même loi (i.e. pour tout $(i,j) \in \mathbb{N}^2$ $X_i \sim X_j$).*
Si les X_i suivent une loi classique, on dira, par exemple, que $(X_n)_{n \in \mathbb{N}}$ est une suite i.i.d. de variables suivant une loi binomiale $B(n,p)$.

Corollaire I.35 (Modélisation d'une répétition infinie d'expériences indépendantes)**.** *Soit P une loi de probabilités discrète sur E. Alors il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}}$ une suite i.i.d. de variables de loi P.*

En appliquant cela à une loi de Bernoulli $B(p)$, on obtient l'existence d'un espace probabilisé modélisant une succession infinie de tirages (éventuellement déséquilibrés) à Pile ou Face.

Corollaire I.36 (Espace probabilisé modélisant des tirages à P/F)**.** *Il existe un espace probabilisé modélisant une succession infinie de tirages à Pile ou Face indépendants, avec une probabilité p d'obtenir Pile à chaque tirage.*

II. Espérance et variance

On étend les définitions vues en première année au cas des variables discrètes. Dans toute cette section, $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé qu'on ne réintroduira pas à chaque énoncé ; sauf mention explicite du contraire, les variables sont définies sur $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

II.1 Espérance d'une variable positive

Définition II.1 (Espérance d'une variable positive)**.** *Soit X une variable aléatoire à valeurs dans $\overline{\mathbb{R}}_+$. L'espérance de X est définie par*

$$\mathbb{E}(X) = \sum_{x\in X(\Omega)} x\mathbb{P}(X = x),$$

où, par convention, $+\infty \times 0 = 0$. L'espérance est toujours bien définie dans $\overline{\mathbb{R}}_+$.

Remarque II.2 (15.2.2)**.** *Si $+\infty \in X(\Omega)$: si $\mathbb{P}(X = +\infty) > 0$, alors $\mathbb{E}(X) = +\infty$; si $\mathbb{P}(X = +\infty) = 0$, alors X a même espérance que sa restriction à $\Omega' = \Omega\backslash\{X = +\infty\}$.*

Plus généralement,

Proposition II.3 (Espérance de variables de même loi)**.** *Si X et Y sont des variables à valeurs dans $\overline{\mathbb{R}}_+$, et $X \sim Y$, alors $\mathbb{E}(X) = \mathbb{E}(Y)$.*

Remarque II.4 (15.2.4)**.** *Une variable aléatoire X positive peut admettre une espérance infinie même si X ne prend que des valeurs finies.*

Exemple II.5 (15.2.5)**.**

$$X(\Omega) = \mathbb{N}^* \text{ et } \forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = \frac{6}{\pi^2 n^2}.$$

Proposition II.6 (Positivité et stricte positivité)**.** *Soit X une variable aléatoire à valeurs dans $\overline{\mathbb{R}}_+$.*
1. $\mathbb{E}(X) \geqslant 0$ 2. $\mathbb{E}(X) = 0 \iff X \sim 0$ (i.e. X est presque sûrement nulle).

Dans le cas de variables à valeurs dans $\mathbb{N}$, on dispose d'une autre description possible de $\mathbb{E}(X)$.

Proposition II.7 (Réexpression de l'espérance lorsque $X(\Omega) \subset \mathbb{N} \cup \{+\infty\}$)**.** *Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors*

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).$$