# Introduction aux probabilités discrètes

Nicolas Schabanel

Décembre 2013

## 1 Espace probabilisé

***Espace des états.*** Nous appellerons *espace des événements* ou *espace des états*, $\Omega$ un ensemble dénombrable (discret). Les éléments $\omega \in \Omega$ sont appellés les *événements élémentaires* ou *états possibles* du système.

***Événements.*** $\wp(\Omega)$, l'ensemble des parties de $\Omega$, est l'ensemble des *événements*. Dire " $A \subseteq \Omega$ se produit" signifie que l'un des événements élémentaires $\omega \in A$ se produit, ou encore que le système est dans un état $\omega \in A$. Ainsi dire que $A = \{\omega_1, \omega_2, \ldots\}$ se produit signifie que le système est dans l'état $\omega_1$ *ou* dans l'état $\omega_2$ *ou*...

Des événements $A_1, A_2, \ldots$ deux-à-deux disjoints sont dits *mutuellement exclusifs* (ils ne peuvent pas se produire en même temps).

L'ensemble des événements $\wp(\Omega)$ est (naturellement) stable par complementation, union et intersection dénombrable. Étant donnés deux événements $A, B \subseteq \Omega$,

- $\overline{A}$ désigne l'événement " $A$ ne se produit pas",
- $A \cup B$ désigne l'événement " $A$ *ou* $B$ se produit ",
- $A \cap B$ désigne l'événement " $A$ *et* $B$ se produisent simultanément",
- $A \setminus B$ désigne l'événement " $A$ se produit mais pas $B$",
- $A \quad B$ désigne l'événement " $A$ *ou* $B$ se produisent mais pas en même temps", ...

En fait, dans la théorie générale des probabilités, on demande uniquement à l'espace des événements d'être stable par complémentation, union et intersection dénombrable. Une telle structure est appelée une *$\sigma$-algèbre*, ou *tribu*.

***Prédicat.*** Étant donné un prédicat (une formule logique) $\varphi : \Omega \to \{\text{Vrai}, \text{Faux}\}$, on désigne par $\{\varphi\}$ l'événement affirmant que $\varphi$ est vrai, i.e., $\{\varphi\} = \{\omega \in \Omega : \varphi(\omega) \text{ est vrai}\}$.

**Exemple 1.1** *Pour le tirage d'un dé, $\varphi =$ " le résultat du tirage est pair ".*

***Espace probabilisé.*** Nous dirons que $\Pr : \Omega \to \mathbb{R}$ est une *mesure de probabilité* sur $\Omega$ si $\Pr \omega \geqslant 0$ pour tout $\omega \in \Omega$, et $\sum_{\omega \in \Omega} \Pr \omega = 1$. On étend Pr à tout $A \subseteq \wp(\Omega)$ en posant $\Pr A = \sum_{\omega \in A} \Pr \omega$. $(\Omega, \wp(\Omega), \Pr)$ est appelé *espace probabilisé*.

**Propriétés élémentaires 1.1**

- *Pour tout $A \subseteq \Omega$, $\Pr \overline{A} = 1 - \Pr A$.*
- *Étant donnés $A_1, A_2, \ldots \subseteq \Omega$ des événements* deux à deux disjoints*,*

$$\Pr A_1 \cup A_2 \cup \ldots = \Pr A_1 + \Pr A_2 + \cdots$$

- *Étant donnés $A_1, A_2, \ldots \subseteq \Omega$ des événements quelconques,*

$$\Pr A_1 \cup A_2 \cup \ldots \leqslant \Pr A_1 + \Pr A_2 + \cdots$$

- *Étant donnés deux événements $A$ et $B$, $\Pr A \cup B = \Pr A + \Pr B - \Pr A \cap B$.*

**Exemple 1.2** *Pour le tirage unique d'un dé à 6 faces,* $\Omega = \{1, 2, 3, 4, 5, 6\}$ *et* $\Pr i = 1/6$ *pour tout* $1 \leqslant i \leqslant 6$*. Alors,*

- $\Pr\{$ *le résultat du tirage est* $\leqslant 3\,\} = \sum_{i:i\leqslant 3} \Pr i = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$,
- $\Pr\{$ *le résultat du tirage est pair* $\} = \sum_{i:\mathrm{pair}(i)} \Pr i = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$,
- $\Pr\{$ *le résultat du tirage est pair et* $\leqslant 3\,\} = \sum_{i:\mathrm{pair}(i)\cap i\leqslant 3} \Pr i = \frac{1}{6}$,
- $\Pr\{$ *le résultat du tirage est pair ou* $\leqslant 3\,\} = \frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}$.

## 2 Variables aléatoires et espérance

Soit $(\Omega, \wp(\Omega), \Pr)$ un espace probabilisé.

***Variable aléatoire.*** Une *variable aléatoire* $X$ sur $\Omega$ est une fonction $X : \Omega \to \mathbb{X}$, où $\mathbb{X}$ est un ensemble quelconque. Pour $x \in \mathbb{X}$, on désigne naturellement par $\{X = x\}$ l'événement associé au prédicat $X = x$, i.e., $\{X = x\} = \{\omega : X(\omega) = x\}$.

***Variable aléatoire uniforme.*** Si $\mathbb{X}$ est un ensemble *fini*, on dit qu'une variable aléatoire $X : \Omega \to \mathbb{X}$ est *uniforme* sur $\mathbb{X}$, si $\Pr\{X = x\} = 1/\#\mathbb{X}$ pour tout $x \in \mathbb{X}$.

Les variables aléatoires uniformes jouent un rôle central en algorithmique d'approximation, tout particulièrement pour les problèmes de dénombrement.

**Exemple 2.1** *Soient* $(G, \cdot)$ *un groupe fini,* $X$ *une variable aléatoire quelconque sur* $G$ *et* $U$ *une variable aléatoire uniforme sur* $G$*, telles que* $X$ *et* $U$ *soient indépendantes, alors* $X \cdot U$ *et* $U \cdot X$ *sont uniformes sur* $G$*. Applications :* $G = \mathbb{Z}/n\mathbb{Z}$ *ou* $G = \mathfrak{S}_n$*.*
*En effet, soit* $g \in G$*.* $\Pr\{X \cdot U = g\} = \sum_{x\in G} \Pr\{X = x \text{ et } x \cdot U = g\} = \sum_{x\in G} \Pr\{X = x \text{ et } U = x^{-1} \cdot g\} = \sum_{x\in G} \Pr\{X = x\} \Pr\{U = x^{-1} \cdot g\} = \frac{\sum_{x\in G} \Pr\{X=x\}}{\#G} = \frac{1}{\#G}$.

***Espérance d'une variable aléatoire.*** L'*espérance* $\mathbb{E}[X]$ d'une variable aléatoire à valeurs réelles $X : \Omega \to \mathbb{R}$ est définie par :

$$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega) \Pr \omega.$$

Remarquez que $\mathbb{E}[X]$ est à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$, et n'est définie que si $\mathbb{E}[X^+]$ et $\mathbb{E}[X^-]$, à valeurs dans $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$, sont cohérentes, où $X^+$ et $X^-$ désignent respectivement les parties positive et négative de $X$, i.e., $X^+ = \max(0, X)$ et $X^- = \max(0, -X)$ (ainsi $X = X^+ - X^-$).

**Fait 2.1** *Si* $X$ *est à valeurs dans* $\mathbb{X} \subseteq \mathbb{R}$ *dénombrable, alors* $\mathbb{E}[X] = \sum_{x\in\mathbb{X}} x \Pr\{X = x\}$*.*

***Preuve.***
$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega) \Pr \omega = \sum_{x\in\mathbb{X}} \sum_{\omega:X(\omega)=x} x \Pr \omega = \sum_{x\in\mathbb{X}} x \sum_{\omega:X(\omega)=x} \Pr \omega = \sum_{x\in\mathbb{X}} x \Pr\{X = x\}$. □

**Exemple 2.2** *Soit* $X$ *le résultat du tirage aléatoire d'un dé à 6 faces. Alors,* $\Omega = \{1, 2, \ldots, 6\}$, $\Pr i = \frac{1}{6}$ *et* $X(i) = i$ *pour tout* $i \in \Omega$*. Ainsi,* $\mathbb{E}[X] = \sum_{i=1}^{6} \frac{i}{6} = 3.5$.

**Exercice 2.1** *Soit* $X$ *une variable aléatoire positive à valeurs dans* $\mathbb{N}$*, démontrez que :*

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} \Pr\{X \geqslant x\}.$$

Dans ce cours, nous nous concentrerons sur l'étude de l'espérance des variables aléatoires. Cependant, en général, ceci n'est pas suffisant car la valeur de l'espérance ne donne que la valeur moyenne du résultat d'un tirage, sans indiquer si cette valeur est représentative du résultat réel. Souvent, il faut évaluer également la *variance* var $X = \mathbb{E}[(X-\mathbb{E}[X])^2]$, pour démontrer que la variable $X$ ne s'écarte pas trop de sa moyenne $\mathbb{E}[X]$ et que son espérance est donc bien représentative.

**Exemple 2.3** *Soit $X$ une variable aléatoire valant $0$ avec probabilité $1-1/\sqrt{n}$ et $n$ avec probabilité $1/\sqrt{n}$. Alors, l'espérance $\mathbb{E}[X] = \sqrt{n}$ n'est pas représentative des valeurs réelles d'un tirage puisqu'elle s'écarte arbitrairement des deux valeurs possibles de $X$ quand $n$ tend vers l'infini.*

## 3 Probabilités conditionnelles et indépendance

Soit $(\Omega, \wp(\Omega), \text{Pr})$ un espace probabilisé.

***Probabilité conditionelle.*** Soient $A, B \subseteq \Omega$ deux événements, tels que $\text{Pr}\, B > 0$. On note $\text{Pr}\, AB$ la *probabilité conditionnelle de $A$ sachant $B$*, définie par :

$$\text{Pr}\{A \mid B\} = \frac{\text{Pr}\, A \cap B}{\text{Pr}\, B}$$

Cette probabilité conditionnelle est simplement la probabilité que $A$ et $B$ se produisent en même temps, renormalisée par le fait que $B$ se produit.

**Fait 3.1** $\text{Pr}_B : \wp(\Omega) \to \mathbb{R}$ *telle que* $\text{Pr}_B\, A = \text{Pr}\{A \mid B\}$ *pour tout* $A \subseteq \Omega$*, est une mesure de probabilité sur* $\Omega$*.*

***Preuve.*** Pour tout $A \subseteq \Omega$, $\text{Pr}_B\, A = \frac{\text{Pr}\, A\cap B}{\text{Pr}\, B} \geqslant 0$, et $\displaystyle\sum_{\omega\in\Omega} \text{Pr}_B\, \omega = \frac{\sum_{\omega\in\Omega} \text{Pr}\,\{\omega\} \cap B}{\text{Pr}\, B} = \frac{\sum_{\omega\in B} \text{Pr}\, \omega}{\text{Pr}\, B} = 1$. □

Remarquez que $\text{Pr}_B$ définit également une mesure de probabilité sur les événements de $(B, \wp(B))$.

***Espérance conditionnelle.*** Étant donnés une variable aléatoire $X$ et un événement $A$ tel que $\text{Pr}\, A > 0$, on appelle *espérance conditionnelle de $X$ sachant $A$* la quantité

$$\mathbb{E}[X \mid A] = \sum_{\omega\in\Omega} X(w)\text{Pr}\{\omega \mid A\} = \frac{1}{\text{Pr}\, A} \sum_{\omega\in A} X(w)\, \text{Pr}\, \omega.$$

**Fait 3.2** *Si $\mathbb{X}$ est dénombrable, alors* $\mathbb{E}[X \mid A] = \sum_{x\in\mathbb{X}} x\, \text{Pr}\{X = x \mid A\}$.

**Exercice 3.1** *Démontrez ce fait.*

**Exemple 3.1** *Considérons un jeu où une personne, prénommée Alice, demande à Bob, le résultat $(I, J)$ du tirage de deux dés à 6 faces et calcule la somme $X = I + J$ des résultats. Dans le premier cas (a), Bob lance deux dés à 6 faces et donne à Alice le résultat des deux dés ; dans le second (b), Bob ne lance qu'un seul dé et donne à Alice deux fois le résultat du tirage. Évaluons dans les deux cas, l'espérance de $X$, la probabilité conditionnelle que $I$ soit pair sachant de $J \leqslant 3$, et l'espérance de $I$ sachant que $J \leqslant 3$.*

(a) *Dans le premier cas, $\Omega = \{1, \ldots, 6\} \times \{1, \ldots, 6\}$ et Pr est uniforme sur $\Omega$, i.e., $\text{Pr}(i, j) = \frac{1}{36}$ pour tout $(i, j) \in \Omega$. La variable aléatoire $X : \Omega \to \mathbb{R}$ est définie par $X(i, j) = i + j$ pour tout $(i, j)$. Alors, $\mathbb{E}[X] = \sum_{i,j} \frac{i+j}{36} = 7$. La probabilité conditionnelle que $I$ soit pair sachant que $J \leqslant 3$ vaut $\text{Pr}\{\text{pair}(I) \mid J \geqslant 3\} = \frac{\text{Pr}\{\text{pair}(I) \text{ et } J\leqslant 3\}}{\text{Pr}\{J\leqslant 3\}}$. Or, $\text{Pr}\{\text{pair}(I) \text{ et } J \leqslant 3\} = \sum_{(i,j):\text{pair}(i) \text{ et } j\leqslant 3} \frac{1}{36} = \frac{3\times 3}{36} = \frac{1}{4}$. L'espérance conditionnelle de $I$ sachant que $J \leqslant 3$ vaut : $\mathbb{E}[I \mid J \leqslant 3] = \frac{\sum_{(i,j):j\leqslant 3} i/36}{\sum_{(i,j):j\leqslant 3} 1/36} = 3.5$.*

*(b) Dans le second cas,* $\Omega = \{1,\ldots,6\}\times\{1,\ldots,6\}$*, mais* Pr *n'est plus uniforme sur* $\Omega$ *et vaut* $\Pr(i,j) = \frac{\delta_{ij}}{6}$*, où* $\delta_{ii} = 1$ *et* $\delta_{ij} = 0$ *pour* $i \neq j$*.* $X$ *est définie de la même façon que précédemment par* $X(i,j) = i+j$ *pour tout* $(i,j) \in \Omega$*, et son espérance vaut* $\mathbb{E}[X] = \sum_{i,j}(i+j)\frac{\delta_{ij}}{6} = 7$*, comme précédemment. La probabilité conditionnelle que* $I$ *soit pair sachant que* $J \leqslant 3$ *vaut* $\Pr\{\text{pair}(I) \mid J \geqslant 3\} = \frac{\Pr\{\text{pair}(I) \text{ et } J \leqslant 3\}}{\Pr\{J \leqslant 3\}}$*. Or,* $\Pr\{\text{pair}(I) \text{ et } J \leqslant 3\} = \sum_{(i,j):\text{pair}(i) \text{ et } j\leqslant 3} \frac{\delta_{ij}}{6} = \frac{1}{6}$*. L'espérance conditionnelle de* $I$ *sachant que* $J \leqslant 3$ *vaut :* $\mathbb{E}[I \mid J \leqslant 3] = \frac{\sum_{(i,j):j\leqslant 3} i\cdot\delta_{ij}/6}{\sum_{(i,j):j\leqslant 3} \delta_{ij}/6} = 2$*.*

*Nous constatons que l'espérance de* $X$ *est inchangée, en revanche les probabilité et espérance conditionnelles de* $I$ *sont modifiées. Dans le premier cas, elles semblent indépendantes de* $J$*, dans le second, le conditionnement sur* $J$ *a un impact. Les définitions suivantes vont expliquer ce phénomène.*

***Indépendance de deux événements.*** Étant donnés deux événements $A, B \subseteq \Omega$, on dit $A$ est *indépendant* de $B$ si $\Pr\{A \mid B\} = \Pr A$.

**Fait 3.3** *$A$ est indépendant de $B \iff B$ est indépendant de $A \iff \Pr A \cap B = \Pr A \times \Pr B$.*

***Preuve.*** Il suffit de démontrer que " $A$ est indépendant de $B \iff \Pr A \cap B = \Pr A \times \Pr B$" car cette dernière proposition est symétrique en $A$ et $B$. Dire que $A$ est indépendant de $B$ est équivalent à $\Pr A = \Pr\{A \mid B\} = \Pr(A \cap B)/\Pr B$, i.e., à $\Pr A \cap B = \Pr A \Pr B$. □

**Exercice 3.2** *Démontrez que si $A$ et $B$ sont indépendants, il en va de même de $A$ et $\overline{B}$.*

***Indépendance de plusieurs événements.*** Des événements $A_1, A_2, \ldots \subseteq \Omega$ sont dits *indépendants (dans leur ensemble)* si pour toute suite finie d'indices $i_1 < \cdots < i_k$,

$$\Pr A_{i_1} \cap \ldots \cap A_{i_k} = \Pr A_{i_1} \times \cdots \times \Pr A_{i_k}.$$

**Fait 3.4** *Soient $A_1, A_2, \ldots \subseteq \Omega$ des événements indépendants, alors pour toute paire de sous-ensembles d'indices finis* disjoints *$I$ et $J$, et tout $k \notin I \cup J$,*

$$\Pr\{A_k \mid \bigcup_{i\in I} A_i \cup \bigcup_{j\in J} \overline{A_j}\} = \Pr A_k.$$

**Exercice 3.3** *Démontrez ce fait.*

Parfois, une forme plus légère d'indépendance suffit à l'analyse d'un algorithme. Nous verrons plus tard que ceci permet de limiter la taille de l'espace de probabilité et donc le nombre de bits aléatoires utilisés dans nos algorithmes.

***Événement* k à k *indépendants.*** Des événements $A_1, A_2, \ldots \subseteq \Omega$ sont dits *$k$ à $k$ indépendants* si pour toute suite d'indices de longueur $k$, $i_1 < \cdots < i_k$,

$$\Pr A_{i_1} \cap \ldots \cap A_{i_k} = \Pr A_{i_1} \times \cdots \times \Pr A_{i_k}.$$

**Exercice 3.4** *Démontrez par un exemple que la proposition suivante est fausse : " des événements $A_1, A_2, \ldots$ $k$ à $k$ indépendants sont $\ell$ à $\ell$ indépendants pour tout $\ell < k$ ".*

***Variables indépendantes.*** Deux variables aléatoires $X : \Omega \to \mathbb{X}$ et $Y\Omega \to \mathbb{Y}$ sont *indépendantes* si pour tout $x \in \mathbb{X}$ et $y \in \mathbb{Y}$,

$$\Pr\{X = x \text{ et } Y = y\} = \Pr\{X = x\} \times \Pr\{Y = y\}$$

De même, des variables aléatoires $X_1, X_2, \ldots$, avec $X_i : \Omega \to \mathbb{X}_i$, sont *indépendantes (dans leur ensemble)* (resp. *$k$ à $k$ indépendantes*) si pour tout $(x_1, x_2, \ldots) \in \mathbb{X}_1 \times \mathbb{X}_2 \times \cdots$, les événements $\{X_1 = x_1\}, \{X_2 = x_2\}, \ldots$ sont indépendants (resp. $k$ à $k$ indépendants).

**Exemple 3.2** *Soient $X$ et $Y$ deux bits aléatoires uniformes et indépendants. Alors $X$, $Y$, $X \oplus Y$ sont des bits aléatoires uniformes $2$ à $2$ indépendants, mais pas indépendants dans leur ensemble.*
*En effet, d'après l'exemple 2.1, $X \oplus Y$ est uniforme. De plus, pour $x, y \in \{0,1\}$, $\Pr\{X = x \text{ et } X \oplus Y = y\} = \Pr\{X = x \text{ et } Y = x \oplus y\} = \Pr\{X = x\} \Pr\{Y = x \oplus y\} = \frac{1}{4} = \Pr\{X = x\} \Pr\{X \oplus Y = y\}$. Mais, $\Pr\{X = 0 \text{ et } Y = 0 \text{ et } X \oplus Y = 1\} = 0$.*

**Exercice 3.5** *Généralisez le résultat de l'exemple 3.2 à des variables aléatoires uniformes indépendantes sur un groupe fini quelconque.*
*Utilisez la même technique pour construire $2^n - 1$ bits uniformes $2$ à $2$ indépendants à partir de $n$ bits aléatoires uniformes indépendants dans leur ensemble. Combien de variables uniformes $2$ à $2$ indépendantes obtenez-vous à partir de $n$ variables uniformes indépendantes dans leur ensemble sur un groupe fini ?*

**Exercice 3.6** *Soient $X : \Omega \to \mathbb{X}$ et $Y : \Omega \to \mathbb{Y}$ des variables aléatoires uniformes. Démontrez que $X$ et $Y$ sont indépendantes ssi $(X, Y)$ est une variable aléatoire uniforme sur $\mathbb{X} \times \mathbb{Y}$.*

**Propriété 3.1** *Si $X$ et $Y$ sont deux variables indépendantes, pour tous prédicats $\varphi$ et $\psi$ :*

$$\Pr\{\varphi(X) \mid \psi(Y)\} = \Pr\{\varphi(X)\} \quad \text{et} \quad \mathbb{E}[X \mid \psi(Y)] = \mathbb{E}[X].$$

***Preuve.*** Pour le premier point, il suffit de montrer que $\varphi(X)$ et $\psi(Y)$ sont des variables (Booléennes) indépendantes :

$$\begin{aligned}
\Pr\{\varphi(X) \text{ et } \psi(Y)\} &= \sum_{\omega : \varphi(X(\omega)) \text{ et } \psi(Y(\omega))} \Pr w \\
&= \sum_{(x,y) : \varphi(x) \text{ et } \psi(y)} \sum_{\omega : X(\omega) = x \text{ et } Y(\omega) = y} \Pr \omega \\
&= \sum_{(x,y) : \varphi(x) \text{ et } \psi(y)} \Pr\{X = x \text{ et } Y = y\} \\
&= \sum_{(x,y) : \varphi(x) \text{ et } \psi(y)} \Pr\{X = x\} \Pr\{Y = y\} \\
&= \sum_{x : \varphi(x)} \Pr\{X = x\} \times \sum_{y : \psi(y)} \Pr\{Y = y\} = \Pr\{\varphi(X)\} \times \Pr\{\psi(Y)\}.
\end{aligned}$$

Pour le second,

$$\begin{aligned}
\mathbb{E}[X \mid \psi(Y)] &= \frac{1}{\Pr\{\psi(Y)\}} \sum_{\omega:\psi(Y(\omega))} X(\omega) \Pr\omega \\
&= \frac{1}{\Pr\{\psi(Y)\}} \sum_{(x,y):\psi(y)} \sum_{\omega:X(\omega)=x \text{ et } Y(\omega)=y} x \Pr\omega \\
&= \frac{1}{\Pr\{\psi(Y)\}} \sum_{(x,y):\psi(y)} x \Pr\{X = x \text{ et } Y = y\} \\
&= \frac{1}{\Pr\{\psi(Y)\}} \sum_{(x,y):\psi(y)} x \Pr\{X = x\} \Pr\{Y = y\} \\
&= \frac{1}{\Pr\{\psi(Y)\}} \sum_{x} x \Pr\{X = x\} \times \sum_{y:\psi(y)} \Pr\{Y = y\} = \mathbb{E}[X].
\end{aligned}$$

□

**Exemple 3.3** *Reprenons l'exemple 3.1. Dans le premier cas (a), $I$ et $J$ sont indépendantes :* $\Pr\{\text{pair}(I) \mid J \leqslant 3\} = \Pr\{\text{pair}(I)\}$. *Dans le second cas (b), $I$ et $J$ sont dépendantes :* $\Pr\{\text{pair}(I) \mid J \leqslant 3\} \neq \Pr\{\text{pair}(I)\}$. *En effet, dans le premier cas* $\Pr\{I = i \text{ et } J = j\} = \frac{1}{36} = \Pr\{I = i\} \times \Pr\{J = j\}$. *Et dans le second cas,* $\Pr\{I = i\} \times \Pr\{J = j\} = \frac{1}{36} \neq \Pr\{I = i \text{ et } J = j\} = \delta_{ij}/6$.

## 4 Propriété fondamentale de l'espérance

Soit $(\Omega, \wp(\Omega), \Pr)$ un espace probabilisé. La propriété suivante, d'apparence anodine, est *extrêmement puissante* et permet de résoudre de nombreux problèmes très simplement lorsqu'elle est employée sur les bonnes variables.

**Propriété 4.1 (LINÉARITÉ DE L'ESPÉRANCE)** *Soient $X$ et $Y$ deux variables aléatoires à valeurs réelles* quelconques *(dépendantes ou non), alors, pour toute constante $\alpha \in \mathbb{R}$,*

$$\mathbb{E}[X + \alpha Y] = \mathbb{E}[X] + \alpha \mathbb{E}[Y].$$

***Preuve.*** $\mathbb{E}[X + \alpha Y] = \sum_{\omega\in\Omega}(X(\omega) + \alpha Y(\omega)) \Pr\omega = \sum_{\omega\in\Omega} X(\omega) \Pr\omega + \alpha \sum_{\omega\in\Omega} Y(\omega) \Pr\omega = \mathbb{E}[X] + \alpha\mathbb{E}[Y]$. □

**Exemple 4.1** *Reprenons l'exemple 3.1. Que les variables $I$ et $J$ soient indépendantes (cas (a)) ou non (cas (b)), l'espérance de $X = I + J$ vaut toujours $7 = \mathbb{E}[I] + \mathbb{E}[J]$.*

**Exemple 4.2 (Les marins bourrés)** *À la suite d'un gueuleton, chacun des 40 marins complétement saouls se couche dans une chambre aléatoire différente choisie parmi leurs 40 chambres. Combien de marins dorment-ils en moyenne dans leurs propres chambres ?*

*Il s'agit d'évaluer le nombre moyen de points fixes dans une permutation aléatoire uniforme de* $\{1, \ldots, 40\}$. *Si on essaye d'évaluer cette espérance directement, les dépendances posent de nombreuses difficultés. Mais introduisons plutôt des* variables indicatrices *$X_i$ qui valent $1$ si le $i$-ème marin dort dans sa chambre et $0$ sinon. On cherche à évaluer l'espérance de la somme $X_1 + \cdots + X_{40}$. Ces variables ne sont bien sûr pas indépendantes : par exemple,* $\Pr\{X_2 = 1 \mid X_1 = 1\} = \frac{1}{39} \neq \Pr\{X_2 = 1\} = \frac{1}{40}$, *mais elles ont la même espérance $\mathbb{E}[X_1] = 1 \times \frac{1}{40} + 0 \times \frac{39}{40} = \frac{1}{40}$ ! Ainsi, l'espérance du nombre de marins qui dorment dans leur chambre vaut tout simplement $40 \times \frac{1}{40} = 1$.*

**Exercice 4.1** *Calculez l'espérance du temps mis pour visiter tous les sommets d'une clique en procédant par une marche aléatoire.*

**Indication.** *Introduisez les variables aléatoires $T_1, \ldots, T_n$, où $T_i$ est le temps mis pour visiter le $i$-ème sommet à partir du moment où on a visité le $(i-1)$-ème. Calculer l'espérance des $T_i$, puis utilisez la linéarité de l'espérance.*

Comme ces exemples le démontrent, la linéarité de l'espérance est une propriété très puissante qui, bien utilisée, permet de simplifier considérablement l'obtention de la valeur moyenne de variables aléatoires. Nous aurons l'occasion de le constater de nouveau à plusieurs reprises dans la suite du cours.

## 5 Théorèmes de grandes déviations

Dans cette section, on cherche à borner la probabilité qu'une variable aléatoire dévie significativement de son espérance, i.e. prenne des valeurs très différentes de son espérance.

**Théorème 5.1 (Inégalité de Markov)** *Soit $X$ une variable aléatoire à valeurs réelles positive ou nulle, alors :*

$$\Pr\{X \geqslant \alpha\mathbb{E}[X]\} \leqslant \frac{1}{\alpha}, \quad \textit{pour tout } \alpha > 0.$$

***Preuve.*** Comme $X$ est à valeurs positives,

$$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega) \Pr\omega \geqslant \sum_{\omega\,:\,X(\omega)\geqslant\alpha\mathbb{E}[X]} \underbrace{X(\omega)}_{\geqslant\alpha\mathbb{E}[X]} \Pr\omega \geqslant \alpha\mathbb{E}(X)\left(\sum_{\omega\,:\,X(\omega)\geqslant\alpha\mathbb{E}[X]} \Pr\omega\right) = \alpha\mathbb{E}[X]\Pr\{X \geqslant \alpha\mathbb{E}[X]\}.$$

Ainsi, soit $\mathbb{E}[X] = 0$ et la propriété est trivialement vraie ; soit $\mathbb{E}[X] > 0$ et on obtient l'inégalité voulue en simplifiant par $\mathbb{E}[X]$ aux deux bouts. □

**Définition 5.1** *On appelle* variance *d'une variable aléatoire $X$ à valeurs réelles la quantité :* $Var(X) = \sigma^2(X) =_{def} \mathbb{E}[(X - \mathbb{E}[X])^2]$.

**Fait 5.1** *Pour toute variable aléatoire $X$ à valeurs réelles :* $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

***Preuve.*** Par linéarité de l'espérance et puisque $\mathbb{E}[X]$ est une valeur constante :
$Var(X) = \mathbb{E}\big[X^2 + 2X\mathbb{E}[X] + \mathbb{E}[X]^2\big] = \mathbb{E}[X^2] - \mathbb{E}\big[2X\mathbb{E}[X]\big] + \mathbb{E}\big[\mathbb{E}[X]^2\big] = \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. □

**Exercice 5.1** *Démontrer qu'une variable aléatoire de variance nulle est une fonction constante égale à son espérance.*

**Théorème 5.2 (Inégalité de Bienaymé-Tchebychev)** *Soit X une variable aléatoire d'espérance $\mathbb{E}[X]$ et de variance finie $\sigma^2$. Pour tout réel $\alpha > 0$,*

$$\Pr\{\,|X - \mathbb{E}[X]| \geqslant \alpha\,\} \leqslant \frac{\sigma^2}{\alpha^2}\,.$$

***Preuve.*** Remarquons que si $\mathbb{E}[Y] = 0$, c'est-à-dire si $X$ est de variance nulle, l'inégalité est trivialement vérifiée. Supposons à présente que $Var(X) = \sigma^2 \neq 0$. Observons que : $\Pr\{|X - \mathbb{E}[X]| \geqslant \alpha\} = \Pr\{(X - \mathbb{E}[X])^2 \geqslant \alpha^2\}$, ainsi en appliquant l'inégalité de Markov à la variable aléatoire positive $Y = (X - \mathbb{E}[X])^2$ telle que $\mathbb{E}[Y] = Var(X) = \sigma^2 > 0$, on obtient :

$$\Pr\{|X - \mathbb{E}[X]| \geqslant \alpha\} = \Pr\left\{Y \geqslant \alpha^2\right\} = \Pr\left\{Y \geqslant \frac{\alpha^2}{\sigma^2}\cdot\mathbb{E}[Y]\right\} \leqslant \frac{\sigma^2}{\alpha^2}$$

□

**Exercice 5.2** *Soient $X_1, \ldots, X_n$ n variables aléatoires* indépendantes *d'espérance $\mu$ et de variance $\sigma^2$. Démontrer que :*

$$\sigma^2(X_1 + \cdots + X_n) = n\sigma^2$$

*En conclure que :*

$$\Pr\{\,|X_1 + \cdots + X_n - n\mu| \geqslant \alpha\sqrt{n}\,\} \leqslant \frac{\sigma^2}{\alpha^2}\,.$$

Lorsque les valeurs des variables aléatoires sont bornées, on peut obtenir un résultat encore plus précis :

**Théorème 5.3 (Inégalité de Hoeffding)** *Soient $X_1, \ldots, X_n$ n variables réelles aléatoires* indépendantes *vérifiant, pour deux suites $a_1, \ldots, a_n$ et $b_1, \ldots, b_n$ de nombres réels tels que $a_k < b_k$, et $(\forall k)$ $\mathrm{Pr}\{a_k \leqslant X_k \leqslant b_k\} = 1$. On pose $S_n = X_1 + X_2 + \cdots + X_n$.*

*Alors, pour tout $t > 0$,*

$$\begin{aligned}
\mathrm{Pr}\left\{S_n - \mathbb{E}[S_n] \geqslant t\right\} &\leqslant \exp\left(-\frac{2\,t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \\
\mathrm{Pr}\left\{S_n - \mathbb{E}[S_n] \leqslant -t\right\} &\leqslant \exp\left(-\frac{2\,t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \\
\mathrm{Pr}\left\{|S_n - \mathbb{E}[S_n]| \geqslant t\right\} &\leqslant 2\exp\left(-\frac{2\,t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).
\end{aligned}$$

**Exercice 5.3** *Soient $X_1, \ldots, X_n$ n variables aléatoires* indépendantes *dont les valeurs sont comprises entre $0$ et $1$, et ayant pour espérance $\mu$, que dire de leur somme ?*

# 6 Couplages probabilistes

Un couplage probabiliste permet de simplifier l'étude d'une variable aléatoire partiellement maitrisée en la ramenant à l'étude d'une autre au comportement parfaitement déterminé.

**Définition 6.1** *Étant donné deux variables aléatoires $X$ et $Y$, un couplage probabiliste $Z = (Z_1, Z_2)$ est une variable aléatoire telle que pour tout $a$ et pour tout $b$ :*

$$\mathrm{Pr}\{Z_1 = a\} = \mathrm{Pr}\{X = a\} \text{ et } \mathrm{Pr}\{Z_2 = b\} = \mathrm{Pr}\{Y = b\}$$

Concrétement, les deux coordonnées d'un couplage probabiliste suivent, quand une seule est considérée, les lois des variables $X$ et $Y$. L'utilité des couplages est d'"aligner" les comportements de deux variables aléatoires pour démontrer des choses sur le comportement de l'une à partir de propriétés de l'autre : par exemple, il est possible de coupler deux variables uniformes sur $\{1, 2\}$ et $\{1, 2, 3, 4\}$ de sorte qu'avec probabilité $1$, les deux variables aient la même parité : ($Z$ a probabilité $1/4$ de valoir $(1, 1)$, $(2, 2)$, $(1, 3)$ ou $(2, 4)$ et $0$ pour les autres paires). Cela démontre que les distributions de la parité de ces deux variables sont identiques

**Définition 6.2 (Dominance stochastique)** *Soient deux variables aléatoires réelles $X$ et $Y$, on dit que $Y$* domine stochastiquement *$X$ s'il existe un couplage probabiliste $Z = (Z_1, Z_2)$ de $X$ et $Y$ tel que :*

$$\mathrm{Pr}\{Z_1 \leqslant Z_2\} = 1$$

*On note parfois $Y \succcurlyeq X$.*

**Exercice 6.1** *Démontrer que si $Y \succcurlyeq X$, alors $\mathbb{E}[Y] \geqslant \mathbb{E}[X]$.*