RÉSUMÉ DE COURS : VARIABLES ALÉATOIRES DISCRÈTES

1 Définitions et premières propriétés

Lorsque l'on intéresse aux phénomènes aléatoires, on est souvent amené à étudier des événements associés aux valeurs d'une fonction. Par exemple, lors d'un jet de deux dés non pipés, on peut considérer l'ensemble des jets dont la somme des résultats vaut 10, c'est-à-dire $f^{-1}(\{10\})$, où f est la fonction qui, à tout jet, associe la somme de ses résultats. De même, si l'on considère les tirages simultanés de trois cartes d'un jeu de 54 cartes, on peut s'intéresser au nombre R de cartes rouges que l'on obtient. Dans ce cas, R est une application de l'univers Ω dans $\mathbb{R}$, et les événements du type "Il y a i cartes rouges tirées" se réécrivent sous la forme $R^{-1}(\{i\})$. Plus généralement, on a la :

Définition 1.1. *Soit* $(\Omega, \mathcal{A})$ *un espace probabilisable. Une variable aléatoire sur* $(\Omega, \mathcal{A})$ *est une application* $X : \Omega \longrightarrow \mathbb{R}$ *telle que, pour tout* $x \in \mathbb{R}$*, l'ensemble* $\{\omega \in \Omega |\ X(\omega) \leq x\}$ *appartient à* $\mathcal{A}$*.*

A noter que cette définition d'une variable aléatoire est équivalente à la définition suivante :

Une variable aléatoire sur $(\Omega, \mathcal{A})$ est une application $X : \Omega \longrightarrow \mathbb{R}$ telle que, pour tout intervalle $I \subset \mathbb{R}$, l'ensemble $X^{-1}(I) = \{\omega \in \Omega |\ X(\omega) \in I\}$ appartient à $\mathcal{A}$.

Définition 1.2. *Soit* X *une variable aléatoire sur un espace probabilisable* $(\Omega, \mathcal{A})$*. Le support de* X*, noté* $X(\Omega)$*, est l'ensemble des valeurs prises par* X*.*

A titre d'exemple, lors du lancer d'un dé équilibré à six faces, la variable aléatoire X égale au résultat du lancer admet pour support $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$. En Théorie des Probabilités, on rappelle qu'on utilise une notation particulière pour les images réciproques d'ensembles par des variables aléatoires. Plus précisément, pour tout réel a et pour toute partie I de $\mathbb{R}$, on pose :

$[X = a] = \{\omega \in \Omega |\ X(\omega) = a\}$	$[X < a] = \{\omega \in \Omega |\ X(\omega) < a\}$,
$[X > a] = \{\omega \in \Omega |\ X(\omega) > a\}$	$[X \leq a] = \{\omega \in \Omega |\ X(\omega) \leq a\}$,
$[X \geq a] = \{\omega \in \Omega |\ X(\omega) \geq a\}$	$[X \in I] = \{\omega \in \Omega |\ X(\omega) \in I\}$.

Théorème 1.3. *Toute combinaison linéaire de variables aléatoires définies sur un même espace probabilisable* $(\Omega, \mathcal{A})$ *est une variable aléatoire sur* $(\Omega, \mathcal{A})$*. De même, tout produit de variables aléatoires définies sur un même espace probabilisable* $(\Omega, \mathcal{A})$ *est une variable aléatoire sur* $(\Omega, \mathcal{A})$*. Enfin, tout maximum et tout minimum de variables aléatoires définies sur un même espace probabilisable* $(\Omega, \mathcal{A})$ *sont des variables aléatoires sur* $(\Omega, \mathcal{A})$*.*

Définition 1.4. *Une variable aléatoire* X *sur un espace probabilisable* $(\Omega, \mathcal{A})$ *est dite discrète si son support* $X(\Omega)$ *est soit fini, soit dénombrable. Si* $X(\Omega)$ *est fini, alors on parle de variable aléatoire finie. Si* $X(\Omega)$ *est dénombrable, alors on parle de variable aléatoire infinie.*

En abrégé, on dira qu'*une variable aléatoire est discrète si son support est au plus dénombrable.* En d'autres termes, ceci revient à écrire qu'il existe une partie $\{x_i\}_{i \in I}$ de $\mathbb{R}$, indexée sur une partie I de $\mathbb{N}$ (finie ou infinie) telle que :

$$\boxed{X(\Omega) = \{x_i\}_{i \in I} \ \text{ et } \ [X = x_i] \text{ est un événement pour tout } i \in I.}$$

On verra plus loin qu'il existe d'autres types de variables aléatoires, qui sont dites *à densité*. A titre d'exemple, lors d'un lancer de deux dés équilibrés, la somme S des résultats du lancer est une variable aléatoire discrète finie. En revanche, lors d'une succession de jeux de pile ou face, le numéro N du premier lancer où l'on obtient un pile est une variable aléatoire discrète infinie (par convention, $N = 0$ si l'on n'obtient jamais de pile). Commençons par voir deux exemples de variables aléatoires discrètes.

Définition 1.5. *Soit* $(\Omega, \mathcal{A})$ *un espace probabilisable, et soit* $a \in \mathbb{R}$*. Alors la variable aléatoire* $X : \Omega \longrightarrow \mathbb{R}$*,* $\omega \longmapsto a$*, est appelée une variable aléatoire constante ou certaine.*

Définition 1.6. *Soit* $(\Omega, \mathcal{A})$ *un espace probabilisable, et soit* $A \in \mathcal{A}$*. Alors la variable aléatoire* $X : \Omega \longrightarrow \mathbb{R}$ *telle que* $X(\omega) = 1$ *si* $\omega \in A$ *et* $X(\omega) = 0$ *sinon, est appelée l'indicatrice de l'événement* A *et notée* $\mathbb{1}_A$*.*

Théorème - définition 1.7. *Soit* X *une variable aléatoire discrète sur un espace probabilisable* $(\Omega, \mathcal{A})$*. Alors la famille* $([X = x])_{x \in X(\Omega)}$ *est un système complet d'événements, que l'on appelle le système complet d'événements associé à* X*.*

2 Loi d'une variable aléatoire

Etant donnée une variable aléatoire X sur un espace probabilisé, on peut alors lui associer sa loi, c'est-à-dire une fonction qui va mesurer la probabilité ou la fréquence qu'a cette variable aléatoire de prendre certaines valeurs. Plus précisément :

Définition 2.1. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé, et soit* X *une variable aléatoire sur* $(\Omega, \mathcal{A}, P)$*. Alors la loi de* X *est la donnée des probabilités* $P([X \in I])$*, où* I *est un intervalle de* $\mathbb{R}$*.*

Dans le cas d'une variable aléatoire discrète, on dispose d'une définition équivalente mais plus simple d'utilisation :

Définition 2.2. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé, et soit* X *une variable aléatoire discrète sur* $(\Omega, \mathcal{A}, P)$*. Alors la loi de* X *est l'application :*

$$P_X : \left\{ \begin{array}{ccc} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P([X = x]) \end{array} \right. .$$

En général, on écrira le support de X sous la forme $X(\Omega) = \{x_i\}_{i \in I}$, où I désigne une partie de $\mathbb{N}$. Bien souvent, dans le cas d'une variable aléatoire finie, le support s'écrira sous la forme $X(\Omega) = \{x_1, ..., x_n\}$, et l'on représentera la loi de X à l'aide d'un tableau formé de deux lignes. Sur la première ligne, on indiquera les éléments de $X(\Omega)$, et sur la deuxième ligne les valeurs correspondantes de P_X :

x_i	x_1	x_2	x_3	...	x_n
$P([X = x_i])$	p_1	p_2	p_3	...	p_n

Théorème 2.3. *Soit* X *une variable aléatoire discrète sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$*, telle que* $X(\Omega) = \{x_i\}_{i \in I}$*, où* I *désigne une partie de* $\mathbb{N}$*. Alors :*

$$\sum_{x \in X(\Omega)} P([X = x]) = \sum_{i \in I} P([X = x_i]) = 1.$$

Théorème 2.4. *Soit* $\{r_i\}_{i \in I}$ *une famille de réels* ≥ 0 *indexée sur une partie* I *de* $\mathbb{N}$*, telle que la série* $\sum r_i$ *converge et* $\sum_{i \in I} r_i = 1$*. Alors il existe un espace probabilisé* $(\Omega, \mathcal{A}, P)$ *et une variable aléatoire réelle discrète* X*, de support* $X(\Omega) = \{x_i\}_{i \in I}$*, tels que* $P([X = x_i]) = r_i$ *pour tout* $i \in I$*.*

On peut considérer ce résultat comme une sorte de "réciproque" du théorème 2.3, puisqu'il nous dit que toute famille au plus dénombrable de réels ≥ 0 dont la somme existe et est égale à 1 peut être considérée comme l'ensemble des valeurs de la loi d'une variable aléatoire discrète. On retiendra surtout qu'*étant donnée une famille* $\{r_i\}_{i \in I}$ *de réels* ≥ 0 *indexée sur une partie* I *de* $\mathbb{N}$*, celle-ci définit la loi de probabilité d'une variable aléatoire discrète si et seulement si la série* $\sum r_i$ *converge et sa somme vaut* 1. A noter que *deux variables aléatoires qui suivent la même loi ne sont pas nécessairement égales.* Par exemple, lors d'une série de n "pile ou face", le nombre X de piles et le nombre Y de faces obtenues ne sont pas égaux. Mais par contre, les lois de X et de Y sont égales.

3 Fonction de répartition d'une variable aléatoire

Définition 3.1. *Soit* X *une variable aléatoire sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$*. Alors la fonction de répartition de* X *est l'application :*

$$F_X : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P([X \leq x]) \end{array} \right. .$$

Proposition 3.2. *Soit* X *une variable aléatoire sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$*. Alors, pour tous réels* a, b *tels que* $a < b$*, on a :*

$$P([a < X \leq b]) = P([X \in]a, b]]) = F_X(b) - F_X(a).$$

Proposition 3.3. *Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Alors sa fonction de répartition F_X est croissante sur $\mathbb{R}$, continue à droite en tout point de $\mathbb{R}$, et de plus :* $\lim_{x\to-\infty} F_X(x) = 0$ *et* $\lim_{x\to+\infty} F_X(x) = 1$.

De façon générale, on montre (et nous admettrons) que, pour tout $x \in \mathbb{R}$:

$$\boxed{\lim_{t\to x^-} F_X(t) = F_X(x) - P([X = x]).}$$

Dès lors, on voit que l'on peut toujours retrouver la loi de X à partir de sa fonction de répartition. En effet, si l'on connaît la fonction de répartition de X, alors les résultats ci-dessus nous permettent de calculer $P([X \in I])$ pour tout intervalle I de $\mathbb{R}$. En particulier, *la fonction de répartition de X caractérise complètement la loi de X*. Plus précisément :

Théorème 3.4. *Deux variables aléatoires sur un même espace probabilisé, dont les fonctions de répartition sont égales, suivent la même loi.*

Dans le cas particulier d'une variable aléatoire discrète X, on dispose des résultats suivants qui permettent de reconstruire la fonction de répartition de X à l'aide de sa loi. Plus précisément :

Proposition 3.5. *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, telle que $X(\Omega) = \{x_i\}_{i\in I}$, où I désigne une partie de $\mathbb{N}$. Alors, pour tout $x \in \mathbb{R}$, on a :*

$$F_X(x) = \sum_{i\in I,\, x_i\leq x} P([X = x_i]).$$

En particulier, si X ne prend que des valeurs entières, alors on obtient la :

Proposition 3.6. *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, telle que $X(\Omega) \subset \mathbb{Z}$. Alors, pour tout $k \in \mathbb{Z}$, on a :*

$$P([X = k]) = F_X(k) - F_X(k - 1).$$

A noter que, si X est une variable aléatoire finie, c'est-à-dire si $X(\Omega) = \{x_1, ..., x_n\}$ avec $x_1 < ... < x_n$, alors F_X est une fonction en escalier croissante (tout comme la partie entière), c'est-à-dire constante sur $[x_k, x_{k+1}[$ pour tout $k \in \{1, ..., n-1\}$, ainsi que sur $]-\infty, x_1[$ et $[x_n, +\infty[$. En particulier, on voit bien que F_X est continue à droite en tout point de $\mathbb{R}$ (comme ceci est annoncé dans la proposition 3.3). A noter enfin qu'en général, on cherche à déterminer directement la loi d'une variable aléatoire discrète X. Mais dans certains cas, il peut être plus facile de calculer d'abord la fonction de répartition de X, puis d'en déduire sa loi (notamment si X est définie à l'aide de maximum et de minimum par exemple).

4 Transformée d'une variable aléatoire discrète

Théorème 4.1. *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, et soit g une application de $X(\Omega)$ dans $\mathbb{R}$. Alors l'application Y de Ω dans $\mathbb{R}$, définie pour tout $\omega \in \Omega$ par $Y(\omega) = g(X(\omega))$, est une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$, notée $g(X)$.*

En Théorie des Probabilités, il est fréquent de considérer des fonctions obtenues en composant (à gauche) une variable aléatoire X avec des applications (comme X^2, $|X|$, e^X, etc). Le résultat ci-dessus nous assure que la composition (à gauche) transforme bien une variable aléatoire en une variable aléatoire. A noter que la loi de Y se déduit de celle de X. Plus précisément, on a le :

Théorème 4.2. *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, et soit g une application de $X(\Omega)$ dans $\mathbb{R}$. Alors la loi de $Y = g(X)$ est donnée pour tout $y \in Y(\Omega)$ par :*

$$P([Y = y]) = \sum_{x\in X(\Omega),\, y=g(x)} P([X = x]).$$

5 Moments d'une variable aléatoire discrète

5.1 Espérance d'une variable aléatoire discrète

Définition 5.1. *Soit X une variable aléatoire discrète finie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, telle que $X(\Omega) = \{x_1, ..., x_n\}$. Alors l'espérance (mathématique) de X est définie par :*

$$E(X) = \sum_{i=1}^{n} x_i P([X = x_i]).$$

En particulier, on voit que $E(1) = 1$. La notion d'espérance se généralise au cas des variables aléatoires discrètes infinies. Plus précisément :

Définition 5.2. *Soit X une variable aléatoire discrète infinie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, telle que $X(\Omega) = \{x_i\}_{i \in I}$, où I est une partie infinie de $\mathbb{N}$. Alors on dit que X admet une espérance si la série $\sum x_i P([X = x_i])$ converge absolument. Dans ce cas, l'espérance de X est définie comme la somme de cette série, c'est-à-dire :*

$$E(X) = \sum_{i \in I} x_i P([X = x_i]).$$

A noter qu'une variable aléatoire finie admet toujours une espérance, ce qui n'est pas forcément le cas d'une variable aléatoire discrète infinie ! De façon générale, on voit que l'espérance de X (si elle existe) est la moyenne des valeurs de X, pondérée par les probabilités qu'a la variable aléatoire X de prendre ces valeurs. Grosso modo, elle correspond à *la valeur moyenne de* X, c'est-à-dire à la valeur que l'on peut raisonnablement s'attendre à obtenir pour X lors d'une expérience aléatoire. *A noter enfin que, sous l'hypothèse de convergence absolue, on voit d'après le cours sur les séries que la définition de l'espérance (comme somme d'une série) ne dépend pas de l'ordre dans lequel on somme les termes.* Dès lors, si $X(\Omega)$ est infini, on dira que X admet une espérance si et seulement si la série $\sum_{x \in X(\Omega)} xP([X = x])$ converge absolument et dans ce cas, son espérance est donnée par :

$$\boxed{E(X) = \sum_{x \in X(\Omega)} xP([X = x]).}$$

A noter que cette écriture fait toujours sens si X est une variable aléatoire finie puisque dans ce cas, il n'y a pas de problème de convergence absolue (vu que la somme est finie).

Théorème 5.3. (Transfert) *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, telle que $X(\Omega) = \{x_i\}_{i \in I}$, où $I \subset \mathbb{N}$, et soit g une application de $X(\Omega)$ dans $\mathbb{R}$.*

1. *Si X est une variable aléatoire finie, alors $g(X)$ admet une espérance et de plus :*

$$E(g(X)) = \sum_{x \in X(\Omega)} g(x)P([X = x]) = \sum_{i \in I} g(x_i)P([X = x_i]).$$

2. *Si X est une variable aléatoire infinie, alors $g(X)$ admet une espérance si et seulement si la série $\sum_{x \in X(\Omega)} g(x)P([X = x])$ converge absolument, et dans ce cas :*

$$E(g(X)) = \sum_{x \in X(\Omega)} g(x)P([X = x]) = \sum_{i \in I} g(x_i)P([X = x_i]).$$

Comme on le voit, la transformée $g(X)$ d'une variable aléatoire finie X admet toujours une espérance, puisqu'elle est elle-même une variable aléatoire finie ! De façon générale, le théorème de transfert est particulièrement utile pour calculer l'espérance (si elle existe) d'une variable aléatoire discrète de la forme $g(X)$, et ce directement à partir de la loi de X, c'est-à-dire sans avoir besoin de calculer la loi de $g(X)$ au préalable. Bien utilisé, il permet de réduire considérablement les calculs.

Théorème 5.4. (Existence de l'espérance par domination) *Soient X, Y deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, telles que $0 \leq |X| \leq Y$ presque sûrement. Si Y admet une espérance, alors X admet aussi une espérance et de plus : $|E(X)| \leq E(Y)$.*

A noter que la condition "$|X| \leq Y$ presque sûrement" signifie que $P([|X| > Y]) = 0$. Dans la pratique, on utilise le théorème d'existence de l'espérance par domination comme suit. Pour montrer qu'une variable aléatoire discrète X admet une espérance, il suffit de majorer $|X|$ par une variable aléatoire discrète plus simple Y, dont on sait qu'elle admet une espérance.

Théorème 5.5. (Linéarité de l'espérance) *Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, et soit $(a, b) \in \mathbb{R}^2$. Si X et Y admet une espérance, alors $aX + bY$ admet aussi une espérance, et de plus :*

$$E(aX + bY) = aE(X) + bE(Y).$$

A noter que la combinaison linéaire $aX + bY$ est bien une variable aléatoire d'après le théorème 1.3. A noter aussi que le théorème 5.5 permet non seulement d'établir l'existence de l'espérance pour une combinaison linéaire de variables aléatoires, mais aussi de la calculer plus facilement en réduisant considérablement les calculs (puisque l'on n'a pas à déterminer la loi de $aX + bY$).

Théorème 5.6. (Positivité de l'espérance) *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Supposons que X soit positive, c'est-à-dire que $X(\omega) \geq 0$ pour tout $\omega \in \Omega$. Si X admet une espérance, alors on a $E(X) \geq 0$, et de plus $E(X) = 0$ si et seulement si X est nulle presque sûrement, c'est-à-dire si $P([X = 0]) = 1$.*

Théorème 5.7. (Croissance de l'espérance) *Soient X, Y deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$, telles que $X \leq Y$ presque sûrement. Si X et Y admettent toutes deux une espérance, alors on a : $E(X) \leq E(Y)$.*

5.2 Variance d'une variable aléatoire discrète

Aux paragraphes précédents, on a vu que l'espérance de X (si elle existe) correspondait à la valeur moyenne de X que l'on pouvait raisonnablement s'attendre à obtenir lors d'une expérience aléatoire. Cependant, on sait qu'en général, cette valeur n'est pas réalisée. Par exemple, si l'on considère un lancer de dé, alors on voit que l'espérance du résultat est égale à $7/2$, alors même que cette valeur n'est jamais atteinte lors d'un lancer. C'est pourquoi nous allons introduire la notion de variance, qui permet de mesurer l'écart entre la valeur réelle obtenue par X lors d'une expérience aléatoire et sa valeur moyenne donnée par l'espérance. Qui plus est, la variance se prête particulièrement bien aux calculs.

Définition 5.8. *Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, et soit $r \in \mathbb{N}^*$. Si X^r admet une espérance, alors son espérance est appelé le moment d'ordre r de X et noté $E(X^r)$ ou $m_r(X)$. Dans ce cas, on dit que X admet un moment d'ordre r.*

A noter que, si X est une variable aléatoire finie de support $X(\Omega) = \{x_1, ..., x_n\}$, alors X^r l'est aussi, et donc X admet un moment d'ordre r pour tout $r \in \mathbb{N}^*$, lequel est donné via le théorème de transfert par :

$$m_r(X) = \sum_{i=1}^{n} x_i^r P([X = x_i]).$$

En outre, si $X(\Omega) = \{x_i\}_{i \in I}$ où I est une partie infinie de $\mathbb{N}$, alors on voit par transfert que X admet un moment d'ordre $r \in \mathbb{N}^*$ si et seulement si la série $\sum x_i^r P([X = x_i])$ converge absolument, et dans ce cas :

$$m_r(X) = \sum_{i \in I} x_i^r P([X = x_i]).$$

De façon générale, on voit que le moment d'ordre r de X (*sous réserve d'existence*) est donné par :

$$\boxed{m_r(X) = \sum_{x \in X(\Omega)} x^r P([X = x]).}$$

Théorème 5.9. *Soit X une variable aléatoire discrète, et soit $r \in \mathbb{N}^*$. Si X admet un moment d'ordre r, alors X admet un moment d'ordre k pour tout entier naturel $k \leq r$.*

En particulier, une telle variable aléatoire admet une espérance puisque, sous réserve d'existence, l'espérance de X n'est ni plus ni moins que son moment d'ordre 1.

Théorème - définition 5.10. *Soit X une variable aléatoire discrète. Si X admet un moment d'ordre 2, alors $X - E(X)$ admet un moment d'ordre 2. Dans ce cas, on dit que X admet une variance, et la variance de X est le moment d'ordre 2 de $X - E(X)$, c'est-à-dire : $V(X) = E\left((X - E(X))^2\right)$.*

Comme précédemment, on voit que, si X est une variable aléatoire finie de support $X(\Omega) = \{x_1, ..., x_n\}$, alors X admet une variance, laquelle est donnée via le théorème de transfert par :

$$V(X) = \sum_{i=1}^{n} (x_i - E(X))^2 P([X = x_i]).$$

En outre, si $X(\Omega) = \{x_i\}_{i \in I}$ où I est une partie infinie de $\mathbb{N}$, et si X admet un moment d'ordre 2, alors on voit par transfert que X admet une variance donnée par :

$$V(X) = \sum_{i \in I} (x_i - E(X))^2 P([X = x_i]).$$

De façon générale, on voit que la variance de X (*sous réserve d'existence*) est donnée par :

$$\boxed{V(X) = \sum_{x \in X(\Omega)} (x - E(X))^2 P([X = x]).}$$

En particulier, $V(X)$ est un réel positif, ce qui justifie la :

Définition 5.11. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors l'écart-type de X est le réel $\sigma(X) = \sqrt{V(X)}$.*

En d'autres termes et sous réserve d'existence, on peut interpréter l'écart-type de X comme la racine de la moyenne des carrés des écarts de la variable aléatoire X avec son espérance. Il fournit ainsi une bonne estimation de l'écart entre la valeur prise par X lors d'une expérience aléatoire et son espérance. De façon plus mathématique, on dira que *la variance de X mesure la dispersion de X (par rapport à son espérance)*. Directement à partir du théorème 5.6, on montre le :

Théorème 5.12. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors $V(X) = 0$ si et seulement si $X = E(X)$ presque sûrement, c'est-à-dire si la probabilité que $X = E(X)$ est égale à 1.*

Théorème 5.13. (Formule de Koenig-Huygens) *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors :*

$$V(X) = E(X^2) - (E(X))^2.$$

La formule de Koenig-Huygens est notamment très utile pour faire des calculs de variance. Dans la pratique, on s'en servira systématiquement pour les calculs.

Théorème 5.14. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors, pour tous réels a et b, la variable aléatoire $aX + b$ admet un moment d'ordre 2, et de plus :*

$$V(aX + b) = a^2 V(X).$$

Définition 5.15. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On dit que X est centrée si $E(X) = 0$, et centrée réduite si $E(X) = 0$ et $\sigma(X) = 1$.*

Théorème - définition 5.16. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2, et de variance non nulle. Alors la variable aléatoire :*

$$X^* = \frac{X - E(X)}{\sigma(X)}$$

est centrée réduite. On dit que X^* est la variable aléatoire centrée réduite associée à X.

6 Les lois usuelles

Nous terminerons ce chapitre par la présentation de quelques lois de probabilité parmi les plus connues. Pour chacune d'entre elles, nous indiquerons dans quels cadres elles apparaissent, et nous en donnerons les caractéristiques principales (espérance, variance, etc).

6.1 Notion de variable certaine

Définition 6.1. *Une variable aléatoire X est certaine si elle ne prend qu'une seule valeur a.*

A noter que, si X est une variable certaine égale à a, alors $F_X(x) = 0$ si $x < a$ et $F_X(x) = 1$ si $x \geq a$. Plus généralement, on dira qu'une variable aléatoire X est *quasi-certaine* s'il existe un réel a tel que $P([X = a]) = 1$. A noter que toute variable certaine est quasi-certaine, mais que la réciproque n'est pas vraie. Plus précisément, il existe des variables quasi-certaines qui ne sont pas certaines. Il s'agit de variables aléatoires qui prennent une valeur a avec probabilité 1, et d'autres valeurs avec probabilité 0.

Théorème 6.2. *Si X est une variable aléatoire certaine avec $X(\Omega) = \{a\}$, alors : $E(X) = a$ et $V(X) = 0$.*

Plus généralement, on voit en vertu du théorème 5.12 qu'*une variable aléatoire discrète X, admettant un moment d'ordre 2, est quasi-certaine si et seulement si $V(X) = 0$.*

6.2 La loi uniforme

Etant donnés deux entiers a, b tels que $a < b$, on désigne par $[\![a, b]\!]$ l'ensemble des entiers compris entre a et b, c'est-à-dire l'ensemble $\{a, a+1, ..., b\}$.

Définition 6.3. *Soit $n \in \mathbb{N}^*$. On dit qu'une variable aléatoire X suit la loi uniforme sur $[\![1, n]\!]$, et on note $X \hookrightarrow \mathcal{U}([\![1, n]\!])$, si $X(\Omega) = [\![1, n]\!]$ et si, pour tout $k \in [\![1, n]\!]$:*

$$P([X = k]) = \frac{1}{n}.$$

A noter que $X(\Omega) = [\![1, n]\!]$. Typiquement, il s'agit de *la loi suivie par une variable aléatoire X qui ne prend que des valeurs entières, et dont toutes les valeurs sont équiprobables.* Par exemple, le résultat d'un lancer de dé suit la loi uniforme sur $[\![1, 6]\!]$. De même, si l'on considère une urne contenant n boules numérotées de 1 à n, alors le résultat du tirage d'une boule de l'urne suit la loi uniforme sur $[\![1, n]\!]$.

Théorème 6.4. *Si $X \hookrightarrow \mathcal{U}([\![1, n]\!])$, alors : $E(X) = \dfrac{n+1}{2}$ et $V(X) = \dfrac{n^2-1}{12}$.*

Définition 6.5. *Soient a et b deux entiers avec $a < b$. On dit qu'une variable aléatoire X suit la loi uniforme sur $[\![a, b]\!]$, et on note $X \hookrightarrow \mathcal{U}([\![a, b]\!])$, si $X(\Omega) = [\![a, b]\!]$ et si, pour tout $k \in [\![a, b]\!]$:*

$$P([X = k]) = \frac{1}{b - a + 1}.$$

Si $X \hookrightarrow \mathcal{U}([\![a, b]\!])$, alors on voit sans peine que la variable $Y = X - a + 1$ suit la loi uniforme sur $[\![1, n]\!]$ avec $n = b - a + 1$. Dans ce cas, les calculs de l'espérance et de la variance nous donnent :

$$\boxed{E(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a)(b-a+2)}{12}.}$$

6.3 La loi de Bernoulli

Définition 6.6. *Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X suit la loi de Bernoulli de paramètre p, et on note $X \hookrightarrow \mathcal{B}(p)$, si $X(\Omega) = \{0, 1\}$ et si $P([X = 1]) = p$.*

A noter que, si X suit la loi de Bernoulli de paramètre p, alors $F_X(x) = 0$ si $x < 0$, $F_X(x) = p$ si $x \in [0, 1[$ et $F_X(x) = 1$ si $x \geq 1$. Typiquement, il s'agit de *la loi suivie par une variable aléatoire X qui ne prend que les valeurs 0 et 1.* Dans ce cas, le paramètre p correspond à la probabilité qu'a X de prendre la valeur 1. Ce genre de loi intervient fréquemment lors d'expériences à deux issues. Par exemple, lors d'un jeu de pile ou face, la variable aléatoire égale à 1 si le résultat est pile et 0 sinon suit la loi de Bernoulli de paramètre 1/2. De même, si l'on extrait une boule d'une urne contenant des boules rouges et bleues, et si p est la proportion de boules rouges dans l'urne, alors la variable aléatoire égale à 1 si l'on tire une boule rouge et 0 sinon suit la loi de Bernoulli de paramètre p. De façon générale, la loi de Bernoulli (de paramètre p) est la loi suivie par l'indicatrice $\mathbb{1}_A$ d'un événement A et dans ce cas, on a $p = P(A)$.

Théorème 6.7. *Si $X \hookrightarrow \mathcal{B}(p)$, alors : $E(X) = p$ et $V(X) = p(1-p)$.*

6.4 La loi binomiale

Définition 6.8. *Soit $n \in \mathbb{N}^*$, soit $p \in]0, 1[$ et posons $q = 1 - p$. On dit qu'une variable aléatoire X suit la loi binomiale de paramètre (n, p), et on note $X \hookrightarrow \mathcal{B}(n, p)$, si $X(\Omega) = [\![0, n]\!]$ et si :*

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

Typiquement, il s'agit de *la loi suivie par le nombre X de succès lors de n répétitions d'une expérience aléatoire à deux issues, où toutes les répétitions se font dans les mêmes conditions.* Plus précisément, considérons une expérience aléatoire à deux issues A et B que l'on répète n fois de suite, telle que l'issue A ait toujours la même probabilité p de se produire à chaque itération. Si X désigne le nombre de fois que l'issue A est réalisée, alors X suit la loi binomiale de paramètre (n, p). Par exemple, si l'on lance n fois de suite une pièce équilibrée, alors le nombre de piles obtenus suit la loi binomiale de paramètre $(n, 1/2)$. De même, si l'on tire n fois de suite et *avec remise* une boule d'une urne contenant des boules rouges et bleues, et si p est la proportion de boules rouges dans l'urne, alors le nombre de boules rouges obtenues lors des n tirages suit la loi binomiale de paramètre (n, p). A noter que, dans ce cas, le nombre de boules bleues obtenues lors des n tirages suit la loi binomiale $\mathcal{B}(n, q)$.

Théorème 6.9. *Si $X \hookrightarrow \mathcal{B}(n, p)$, alors : $E(X) = np$ et $V(X) = npq$.*

6.5 La loi géométrique

Définition 6.10. *Soit $p \in]0,1[$, et posons $q = 1 - p$. On dit qu'une variable aléatoire X suit la loi géométrique de paramètre p, et on note $X \hookrightarrow \mathcal{G}(p)$, si $X(\Omega) = \mathbb{N}^*$ et si :*

$$\forall k \in \mathbb{N}^*, \quad P([X = k]) = pq^{k-1}.$$

Typiquement, il s'agit de *la loi suivie par le temps d'attente X du premier succès lors d'une succession d'expériences à deux issues, où toutes les expériences se font dans les mêmes conditions.* Plus précisément, considérons une expérience aléatoire à deux issues A et B que l'on répète indéfiniment, telle que l'issue A ait toujours la même probabilité p de se produire à chaque itération. Si l'on numérote chacune de ces répétitions, et si X désigne le plus petit numéro pour lequel l'issue A est réalisée, alors X suit la loi géométrique de paramètre p. Par exemple, si l'on lance successivement une pièce équilibrée, alors le temps d'attente du premier pile suit la loi géométrique de paramètre $1/2$.

Théorème 6.11. *Si $X \hookrightarrow \mathcal{G}(p)$, alors : $E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$.*

6.6 La loi de Poisson

Définition 6.12. *Soit λ un réel > 0. On dit qu'une variable aléatoire X suit la loi de Poisson de paramètre λ, et on note $X \hookrightarrow \mathcal{P}(\lambda)$, si $X(\Omega) = \mathbb{N}$ et si :*

$$\forall k \in \mathbb{N}, \quad P([X = k]) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Ici, on n'a pas de modèle simple pour décrire la loi de Poisson, que l'on appelle quelquefois *la loi des événements rares.* C'est pourquoi *on indique toujours dans un sujet au préalable qu'une variable aléatoire suit une loi de Poisson si besoin !* En général, la loi de Poisson apparaît dans des phénomènes de comptage sur une période de temps donné. Plus précisément, considérons un phénomène aléatoire qui peut se répéter de façon ponctuelle au cours du temps. Pour tout $t \geq 0$, on désigne par N_t le nombre d'occurrences de ce phénomène sur l'intervalle de temps $[0, t]$. On dit que la famille de variables aléatoires $\{N_t\}_{t \geq 0}$ définit un *processus de Poisson* si elle vérifie les conditions suivantes :

- le processus est *sans mémoire*, c'est-à-dire que les nombres d'occurrences du phénomène sur des intervalles de temps disjoints sont indépendants,
- le processus est *homogène*, c'est-à-dire que le nombre d'occurrence sur un intervalle de temps donné ne dépend que de la longueur de l'intervalle (et non du début ou de la fin).

Dans ce cas, on montre et nous admettrons qu'il existe un réel $\lambda > 0$ telle que : $\forall t \in \mathbb{R}_+^*$, $N_t \hookrightarrow \mathcal{P}(t\lambda)$. En particulier, la variable aléatoire N_1 suit la loi de Poisson de paramètre λ. A titre d'exemple, on utilise fréquemment la loi de Poisson pour modéliser le nombre de personnes qui vont à la poste sur un laps de temps donné, le nombre de voitures arrivant à un péage sur une journée, le nombre d'appels à un standard téléphonique sur une période donnée, etc.

Théorème 6.13. *Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors : $E(X) = \lambda$ et $V(X) = \lambda$.*