Chapitre 34

Variables aléatoires sur un espace probabilisé fini

Dans tout ce chapitre, on fixe un espace probabilisé fini (Ω, P).

1 Variables aléatoires

1.1 Définitions

Définition 1.1 (Variable aléatoire)

1. Une variable aléatoire sur Ω est une application définie sur Ω.
2. Une variable aléatoire réelle sur Ω est une application définie sur Ω à valeurs dans $\mathbb{R}$.

Exemples.

1. Une urne contient une boule blanche et une boule noire. On tire avec remise deux boules, et on note X le nombre de boules blanches obtenues : X est une variable aléatoire réelle sur l'univers des couples de boules.

2. On lance deux dés, et on note X la somme des points obtenus : c'est une variable aléatoire réelle sur l'univers $[\![1, 6]\!]^2$.

Remarque.
L'univers Ω étant fini, si X est une variable aléatoire sur Ω, son image $X(\Omega)$ est également un ensemble fini. On appelle "univers image" l'ensemble $X(\Omega)$.

Définition 1.2 (Variable aléatoire constante)

Une variable aléatoire constante (ou certaine) sur l'univers Ω est une fonction constante sur Ω.

Exemple.
On tire 5 cartes dans un jeu de 52 cartes. Le nombre X de cartes obtenues est une variable aléatoire constante sur l'univers des parties à 5 éléments de l'ensemble des cartes.

Définition 1.3 (Variable indicatrice d'un événement)

Soit A un événement de Ω. La variable indicatrice de A est la variable aléatoire réelle

$$\mathbb{1}_A : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases} \end{array}$$

Exemple.

On lance un dé une fois. Sur l'univers $\Omega = [\![1, 6]\!]$, on considère la variable aléatoire X qui donne le reste modulo 2 du chiffre obtenu. C'est la variable aléatoire indicatrice de l'événement $\{1, 3, 5\}$ (le chiffre est impair).

Remarques.

1. L'ensemble des variables aléatoires réelles est l'ensemble $\mathbb{R}^\Omega$, qui est un $\mathbb{R}$-espace vectoriel pour l'addition point par point, et la multiplication par les scalaires point par point. Les combinaisons linéaires de variables aléatoires réelles sont donc des variables aléatoires réelles.
2. De même, le produit de deux variables aléatoires réelles est une variable aléatoire réelle.
3. Le minimum et le maximum de deux variables aléatoires réelles en est une aussi.

Proposition 1.4 (Image d'une variable aléatoire)

Soit X une variable aléatoire sur Ω, et f une fonction définie sur $X(\Omega)$. Alors $f \circ X$ est une variable aléatoire sur Ω, notée $f(X)$.

Démonstration.

$f \circ X$ est une application définie sur Ω. ∎

Remarque.

La notation $u(X)$ est cohérente avec le vocabulaire "variable" aléatoire.

Exemple.

Si X est une variable aléatoire réelle, X^2 l'est également, et si X ne s'annule pas, $1/X$ est aussi une variable aléatoire.

1.2 Notations

Soit X une variable aléatoire sur Ω, à valeurs dans un ensemble E. On considère également une probabilité P sur Ω, et $A \in \mathcal{P}(E)$.

1. L'événement $X^{-1}(A) = \{\omega \in \Omega,\ X(\omega) \in A\} \subset \Omega$ se note $\{X \in A\}$ ou $(X \in A)$.
2. Si X est une variable aléatoire réelle, on note
   - $(X = x) = \{X = x\} = \{\omega \in \Omega,\ X(\omega) = x\} = X^{-1}(\{x\})$.
   - $(X \leqslant x) = \{X \leqslant x\} = \{\omega \in \Omega,\ X(\omega) \leqslant x\} = X^{-1}(]-\infty, x])$.
   - $(X \geqslant x) = \{X \geqslant x\} = \{\omega \in \Omega,\ X(\omega) \geqslant x\} = X^{-1}([x, +\infty[)$.

— $(X < x) = \{X < x\} = \{\omega \in \Omega,\ X(\omega) < x\} = X^{-1}(]-\infty, x[)$.
— $(X > x) = \{X > x\} = \{\omega \in \Omega,\ X(\omega) > x\} = X^{-1}(]x, +\infty[)$.

(Ce sont des événements.)

3. On note $P(X \in A) = P(\{X \in A\})$ et $P(X = x) = P(\{X = x\})$.

Exemples.

1. On a $(X = x) = \emptyset$ si et seulement si $x \notin X(\Omega)$.
2. On a $(X \in A \cup B) = (X \in A) \cup (X \cup B)$.
3. On lance deux dés et on note X la somme des chiffres obtenus. Alors
$$(X \geqslant 10) = \{(6,4); (6,5); (6,6); (5,5); (5;6); (4,6)\}.$$

Proposition 1.5

Soit $X : \Omega \longrightarrow E$ une variable aléatoire sur Ω, et $A, B \subset E$. Alors
$$(X \in A) \cup (X \in B) = (X \in A \cup B), \qquad (X \in A) \cap (X \in B) = (X \in A \cap B).$$

Proposition 1.6

Soit $X : \Omega \longrightarrow E$ une variable aléatoire sur Ω, et $A \subset E$. Alors
$$(X \in A) = (X \in A \cap X(\Omega)) = \bigcup_{x \in A \cap X(\Omega)} (X = x).$$

Démonstration.

La première égalité découle de la définition de l'image réciproque de A par X. La deuxième découle de la compatibilité des images réciproques avec l'union : on a
$$A \cap X(\Omega) = \bigcup_{x \in A \cap X(\Omega)} \{x\},$$
donc
$$(X \in A \cap X(\Omega)) = X^{-1}\left(\bigcup_{x \in A \cap X(\Omega)} \{x\}\right) = \bigcup_{x \in A \cap X(\Omega)} X^{-1}(\{x\}) = \bigcup_{x \in A \cap X(\Omega)} (X = x).$$

■

1.3 Loi d'une variable aléatoire

On rappelle qu'on a fixé une probabilité P sur Ω.

Théorème 1.7 (Loi de X)

Soit X une variable aléatoire sur Ω. L'application
$$\begin{array}{ccc} \mathcal{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & P(X \in A) \end{array}$$
est une probabilité sur $X(\Omega)$, appelée loi de X, et notée P_X.

Démonstration.

La fonction P_X est définie sur $\mathcal{P}(X(\Omega)$, et à valeurs dans $[0,1]$. De plus,

$$P_X(X(\Omega)) = P(X \in X(\Omega)) = P(\Omega) = 1,$$

car P est une probabilité. Enfin, si $A, B \in \mathcal{P}(X(\Omega))$ sont disjoints, montrons que $P_X(A \cup B) = P_X(A) + P_X(B)$. Mais on a $(X \in A) \cap (X \in B) = \emptyset$ (car si $x \in (X \in A) \cap (X \cap B)$, alors $X(x) \in A \cap B = \emptyset$), donc

$$P_x(A \cup B) = P(X \in A \cup B) = P\Big((X \in A) \cup (X \in B)\Big) = P(X \in A) + P(X \in B) = P_X(A) + P_X(B).$$

■

Remarques.

1. Quand on s'interesse à la loi de X, on a en fait un nouvel univers, qui est $X(\Omega)$.
2. La loi de X dépend bien entendu de la probabilité P sur Ω.
3. La loi de X est la probabilité qu'un événement ait pour résultat (par X) un élément de A.
4. Attention : deux variables aléatoires différentes peuvent avoir même loi. Par exemple, si on lance un dé, et on note X la variable qui donne 0 si le chiffre est pair, et 1 sinon, et Y la v.a.r. qui fait l'inverse (0 si impair), alors $X \neq Y$, mais les lois sont les mêmes puisqu'on a à chaque fois une probabilité $1/2$ d'avoir 0 ou 1.

Proposition 1.8 (Système complet d'événements associé à X)

Soit X une variable aléatoire sur Ω. La famille $(\{X = x\})_{x \in X(\Omega)}$ est un système complet d'événements de Ω, appelé système complet d'événements associé à X. En particulier,

$$\sum_{x \in X(\Omega)} P(X = x) = 1.$$

Démonstration.

Si $x \neq y$ et $x, y \in X(\Omega)$, on a bien $(X = x) \cap (X = y) = \emptyset$. En effet, si $\omega \in (X = x)$, par définition, $X(\omega) = x \neq y$, donc $\omega \notin (X = y)$.

De plus, si $\omega \in \Omega$, alors $\omega \in (X = X(\omega))$, donc

$$\Omega = \bigcup_{x \in X(\Omega)} (X = x),$$

ce qui prouve qu'on a bien un système complet d'événements. ■

Remarque.
On peut aussi utiliser des notations différentes. Comme l'ensemble $X(\Omega)$ est fini, on peut l'écrire sous la forme $X(\Omega) = \{x_1, \ldots, x_n\}$ ($n \in \mathbb{N}^*$), et $((X = x_i))_{1 \leqslant i \leqslant n}$ est un système complet d'événements, et on a $\displaystyle\sum_{i=1}^{n} P(X = x_i) = 1$.

Proposition 1.9

Soit X une variable aléatoire sur Ω. Pour tout $A \subset X(\Omega)$, on a

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

Autrement dit, la loi de X est entièrement déterminée par la donnée des $\Big(P(X = x)\Big)_{x \in X(\Omega)}$.

Démonstration.

Soit $A \subset X(\Omega)$. On sait que la famille $(X = x)_{x \in X(\Omega)}$ est un sce, donc

$$P(X \in A) = \sum_{x \in X(\Omega)} P((X \in A) \cap (X = x)) = \sum_{x \in A} P(X = x).$$

■

Remarques.

1. Reprenons les notations $X(\Omega) = \{x_1, \ldots, x_n\}$. La loi de X est donc entièrement déterminée par la donnée de $(P(X = x_i))_{1 \leqslant i \leqslant n}$.
2. Attention : on parle d'événements élémentaires de $X(\Omega)$. L'ensemble $(X = x)$ n'est pas un événement élémentaire de Ω.

Méthode 1.10

Pour déterminer la loi d'une variable aléatoire X, on détermine $X(\Omega)$, et pour tout $x \in X(\Omega)$, on calcule $P(X = x)$.

Pour cela, on peut remarquer que $P(X = x) = \displaystyle\sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P(\{\omega\})$.

Exemple.

On considère un jeu où on lance deux dés, et on gagne en euros la somme des chiffres si elle est impaire, et on la perd si elle est paire. Soit G la v.a.r. égale au gain associé au jeu. Alors $\Omega = [\![1, 6]\!]^2$,

$$G(\Omega) = \{-2, -4, -6, -8, -10, -12, 3, 5, 7, 9, 11\}.$$

La loi de G est donnée par

x	-2	-4	-6	-8	-10	-12	3	5	7	9	11
$P(G = x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

Proposition 1.11 (Loi de $f(X)$)

Soit X une variable aléatoire sur Ω, et f une application définie sur $X(\Omega)$. La loi de $f(X)$ est donnée par :

$$\forall\, y \in\ f(X)(\Omega),\ P(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} P(X = x),$$

ou encore, si $X(\Omega) = \{x_1, \ldots, x_n\}$,

$$\forall\, y \in\ f(X)(\Omega),\ P(f(X) = y) = \sum_{\substack{i=1 \\ f(x_i) = y}}^{n} P(X = x_i).$$

Démonstration.

Remarquons d'abord qu'en vertu de la proposition 1.9, la loi de $f(X)$ est déterminée par les $P(f(X) = y)$ pour $y \in f(X)(\Omega)$.

Or, $(f(X) = y) = (X \in f^{-1}(\{y\}))$, car si $\omega \in \Omega$, $f(X)(\omega) = y$ si et seulement si $X(\omega) \in f^{-1}(\{y\})$. Mais par 1.9, on a

$$P(X \in f^{-1}(\{y\})) = \sum_{x \in f^{-1}(\{y\})} P(X = x),$$

ce qui donne la formule annoncée. ■

Exemples.

1. Loi de $aX + b$: si $X(\Omega) = \{x_1, \ldots, x_n\}$, alors

$$P(aX + b = ax_i + b) = P(X = x_i).$$

2. Loi de X^2 : si $X(\Omega) = [\![-n, n]\!]$, alors $X^2(\Omega) = \left\{k^2,\ k \in [\![0, n]\!]\right\}$, et on a, pour $y \in\ X^2(\Omega)$, $y = k^2 > 0$, on a

$$P(X^2 = y) = P(X^2 = k^2) = P(X = k) + P(X = -k),$$

et $P(X^2 = 0) = P(X = 0)$.

3. Soit X une v.a.r. telle que $X(\Omega) = \{-2, -1, 0, 1\}$, et telle que

$$P(X = -2) = P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{4}.$$

Donnons la loi de X^2. On a $X^2(\Omega) = \{0, 1, 4\}$, et

$$\begin{aligned} P(X^2 = 0) &= P(X = 0) = \frac{1}{4}, \\ P(X^2 = 1) &= P(X = -1) + P(X = 1) = \frac{1}{2}, \\ P(X^2 = 4) &= P(X = -2) = \frac{1}{4}. \end{aligned}$$

Remarque.
Il n'y a pas de formule simple donnant la loi de $X + Y$, XY, ou plus généralement $u(X)$. Il faut la plupart du temps tout refaire avec la proposition 1.11.

2 Lois usuelles

2.1 Loi uniforme

Définition 2.1 (Loi uniforme)

1. Soit $n \in \mathbb{N}^*$. Une variable aléatoire X sur Ω suit la loi uniforme sur $[\![1, n]\!]$ si $X(\Omega) = [\![1, n]\!]$, et si

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

2. Soient $a, b \in \mathbb{Z}$ avec $a \leqslant b$. Une variable aléatoire X sur Ω suit la loi uniforme sur $[\![a, b]\!]$ si $X(\Omega) = [\![a, b]\!]$, et si

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

Remarques.

1. Une telle loi traduit le fait qu'on choisit au hasard entre n possibilités.
2. Si X suit la loi uniforme sur $[\![a, b]\!]$, alors $Y = X - a + 1$ suit la loi uniforme sur $[\![1, n]\!]$, où $n = b - a + 1$.
3. On note en général $X \sim \mathcal{U}([\![a, b]\!])$, ou $X \sim \mathcal{U}(n)$ dans le cas particulier de $[\![1, n]\!]$.

Exemple.

On lance un dé, on note X le chiffre obtenu. Alors $X \sim \mathcal{U}([\![1, 6]\!])$.

2.2 Loi de Bernoulli

Définition 2.2 (Loi de Bernoulli)

Soit $p \in [0, 1]$. Une variable aléatoire X suit la loi de Bernoulli de paramètre p si $X(\Omega) \subset \{0, 1\}$, et si $P(X = 1) = p$.

Remarques.

1. Une telle loi modélise une épreuve "succès-échec" d'un événement.
2. On note $X \sim \mathcal{B}(p)$, et on dit parfois que X est une variable de Bernoulli.
3. Si $p \in]0, 1[$, alors $X(\Omega) = \{0, 1\}$.

Proposition 2.3

Soit $p \in [0, 1]$ et X une variable de Bernouilli de paramètre p. Alors $P(X = 0) = 1 - p$.

Démonstration.

L'événement contraire de $(X = 1)$ est $(X = 0)$, donc $P(X = 0) = 1 - P(X = 1) = 1 - p$. ■

Exemple.

On lance un dé à 5 faces (numérotées de 1 à 5), et on note $X = 1$ si le chiffre est pair. Alors $X \sim \mathcal{B}\left(\frac{2}{5}\right)$.

Proposition 2.4 (Indicatrice d'un événement)

1. La variable indicatrice d'un événement A est une variable de Bernoulli, de paramètre $P(A)$.
2. Réciproquement, une variable de bernoulli X est la variable indicatrice de l'événement $(X = 1)$.

Démonstration.

Par définition, si A est un événement, $\mathbb{1}_A$ est à valeurs dans $\{0, 1\}$, et $P(\mathbb{1}_A = 1) = P(A)$.

Réciproquement, si X est une variable de Bernoulli, c'est la variable indicatrice de $A = (X = 1)$. ■

Proposition 2.5

Soient X et Y des variables de Bernoulli.

1. On a $X^2 = X$.

2. XY est une variable de Bernoulli.

Démonstration.

Puisque X ne prend que les valeurs 0 et 1, et que $0^2 = 0$, et $1^1 = 1$, on a $X^2 = X$.

Puis XY ne peut prendre que 0 ou 1 comme valeur, donc XY est bien une variable de Bernoulli. ■

Remarque.

Attention : dans le cas du produit de deux variables de Bernoulli, le paramètre n'est pas le produit des paramètres de X et Y, *cf.* la notion de variables aléatoires indépendantes au paragraphe 6.

2.3 Loi binomiale

Définition 2.6 (Loi binomiale)

Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. Une variable aléatoire sur Ω suit la loi binomiale de paramètre (n, p) si

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

Remarques.

1. On note $X \sim \mathcal{B}(n, p)$.
2. Si $p \in \,]0, 1[$, alors $X(\Omega) = [\![0, n]\!]$.
3. Une variable binomiale de paramètre $(1, p)$ est une variable de Bernoulli de paramètre p : $\mathcal{B}(1, p) = \mathcal{B}(p)$.

Théorème 2.7 (Épreuves de Bernoulli indépendantes)

Soient $n \in \mathbb{N}$ et $(E_k)_{1 \leqslant k \leqslant n}$ des épreuves de Bernoulli indépendantes de même paramètre $p \in [0,1]$. Soit X le nombre de succès. Alors $X \sim \mathcal{B}(n, p)$.

Démonstration.

On a bien sûr un nombre de succès qui est compris entre 0 et n, donc $X(\Omega) \subset [\![1, n]\!]$ (l'univers est $\Omega = \{0, 1\}^n$).

Pour $i \in [\![1, n]\!]$, notons A_i l'événement "la $i$ème épreuve est un succès" (en terme de notre univers, c'est l'union de tous les événements dont la composante i vaut 1). Puisque les épreuves sont indépendantes, les événements $(A_i)_{1 \leqslant i \leqslant n}$ sont indépendants.

Déterminons alors $P(X = k)$ pour $k \in [\![0, n]\!]$. Pour cela, on considère k entiers $1 \leqslant i_1 < \cdots < i_k \leqslant n$, et notons $j_1, \ldots, j_{n-k} = [\![1, n]\!] \setminus \{i_1, \ldots i_k\}$. Considérons l'événement "succès aux épreuves $i_1, \ldots, i_k$ et échec ailleurs", qui est

$$A_{i_1} \cap \cdots A_{i_k} \cap \overline{A_{j_1}} \cap \cdots \overline{A_{j_{n-k}}}.$$

Par indépendance de ces événement, on a

$$P\Big(A_{i_1} \cap \cdots A_{i_k} \cap \overline{A_{j_1}} \cap \cdots \overline{A_{j_{n-k}}}\Big) = P(A_{i_1}) \cdots P(A_{i_k}) P(\overline{A_{j_1}} \cdots P(\overline{A_{j_{n-k}}}) = p^k (1-p)^{n-k}.$$

Or,

$$(X = k) = \bigcup_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \Big(A_{i_1} \cap \cdots A_{i_k} \cap \overline{A_{j_1}} \cap \cdots \overline{A_{j_{n-k}}}\Big),$$

et ces événements sont deux à deux incompatibles, donc

$$P(X = k) = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} p^k (1-p)^{n-k}.$$

Or, c'est une somme à $\binom{n}{k}$ termes (choix du k-uplet à ranger dans l'ordre strictement croissant, donc choix d'une k-combinaison). On en déduit que

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

■

Exemples.

1. On lance n fois un dé à 6 faces, et on note X le nombre de 3 obtenus : $X \sim \mathcal{B}(n, \frac{1}{6})$.

2. On tire avec remise successivement n boules dans une urne contenant une proportion de p boules blanches, et on note X le nombre de boules blanches obtenues. Alors $X \sim \mathcal{B}(n, p)$.

Remarque.

La plupart des lois de variable aléatoire n'ont pas de nom. Il faudra alors donner les probabilités de chacun des événements du s.c.e. associé à la v.a. pour donner la loi.

Voici un exemple. Un jeu payant (5 euros) consiste à lancer une pièce de monnaie deux fois de suites.

Si la pièce donne deux fois le même résultats, on gagne 6 euros.

—Si pile sort en premier, puis face, on perd la mise.

—Si face sort en premier, puis pile, on gagne 7 euros.

Quelle est la loi du gain G (mise comprise) ? On peut modéliser la situation en posant $\Omega = \{P, F\}^2$ et en utilisant la probabilité uniforme sur $\mathcal{P}(\Omega)$. On a $G(\Omega) = \{-5, 1, 2\}$.

Examinons la loi de G :

On a $(X = 1) = \{(P, P), (F, F)\}$, donc $P(X = 1) = \frac{2}{4} = \frac{1}{2}$.

—Puis $(X = -5) = \{(P, F)\}$ donc $P(X = -5) = \frac{1}{4}$.

—Enfin, $(X = 2) = \{(F, P)\}$ donc $P(X = 2) = \frac{1}{4}$.

Pour terminer, la probabilité de gagner est

$$P(X > 0) = P((X = 2) \cup (X = 1)) = P(X = 2) + P(X = 1) = \frac{3}{4}.$$

3 Espérance

On fixe un espace probabilisé (Ω, P).

3.1 Définitions

Définition 3.1 (Espérance)

L'espérance d'une variable aléatoire réelle X est le réel

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

ou encore, si $X(\Omega) = \{x_1, \ldots, x_n\}$,

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

Remarques.

1. C'est la moyenne coefficienté, le gain moyen d'un jeu. Mais attention, ce n'est pas la valeur qu'on obtient en répétant suffisament de fois une épreuve.
2. Deux v.a.r. ayant même loi ont même espérance.

Exemple.

Un jeu payant (5 euros) consiste à lancer une pièce de monnaie deux fois de suites.
Si la pièce donne deux fois le même résultat, on gagne 6 euros.
—Si pile sort en premier, puis face, on perd la mise.
—Si face sort en premier, puis pile, on gagne 7 euros.
On a vu que la loi du gain G est $(G(\Omega) = \{-5, 1, 2\})$:

$$P(X = 1) = \frac{1}{2}, \qquad P(X = -5) = \frac{1}{4}, \qquad P(X = 2) = \frac{1}{4}.$$

et la probabilité de gagner est $\frac{3}{4}$. L'espérance de gain est ici

$$E(G) = \frac{1}{2} - 5 \times \frac{1}{4} + 2 \times \frac{1}{4} = -\frac{1}{4}.$$

À terme, même si on a plus de chance de gagner que de perdre, on va perdre de l'argent ! Il ne faut pas jouer.

Proposition 3.2

Soit X une variable aléatoire réelle sur (Ω, P). Alors

$$E(X) = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega).$$

Démonstration.

Soit $x \in X(\Omega)$. Alors

$$(X = x) = \bigcup_{\omega \in (X=x)} \{\omega\},$$

qui est une union d'événements deux à deux incompatibles, donc

$$xP(X = x) = x \sum_{\omega \in (X=x)} P(\{\omega\}) = \sum_{\omega \in (X=x)} xP(\{\omega\}) = \sum_{\omega \in (X=x)} P(\{\omega\})X(\omega).$$

On en déduit que

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

où la dernière égalité découle du fait que

$$\bigcup_{x \in X(\Omega)} \bigcup_{\omega \in (X=x)} \{\omega\} = \Omega,$$

et que c'est une union d'ensembles deux à deux disjoints. ■

Définition 3.3 (Variable aléatoire centrée)

Une variable aléatoire réelle centrée est une variable aléatoire d'espérance nulle.

Exemple.

Si on reprend le jeu précédent avec 8 euros au lieu de 7 (et $X(\Omega) = \{-5, 1, 3\}$), on obtient une v.a.r. centrée.

3.2 Espérance des lois usuelles

Proposition 3.4 (Espérance d'une variable indicatrice)

Soit A un événement. Alors $E(\mathbb{1}_A) = P(A)$.

Démonstration.

L'univers image est ici $\{0, 1\}$. On a donc $E(\mathbb{1}_A) = 0 \times P(\overline{A}) + 1 \times P(A) = P(A)$. ■

Proposition 3.5 (Espérance d'une variable aléatoire certaine)

Soit $a \in \mathbb{R}$. L'espérance de la variable aléatoire certaine égale à a est a.

Démonstration.

L'univers image est ici $\{a\}$, et la loi est donnée par $P(X = a) = 1$. ■

Proposition 3.6 (Espérance d'une loi uniforme)

Soit X une variable aléatoire réelle.

1. Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, alors $E(X) = \frac{n+1}{2}$.

2. Si $X \sim \mathcal{U}(\llbracket 0, n \rrbracket)$, alors $E(X) = \frac{n}{2}$.

3. Si $X \sim \mathcal{U}(\llbracket a, b \rrbracket)$, alors $E(X) = \frac{a+b}{2}$.

Démonstration.

On traite le troisième point. Les deux premiers en découle aussitôt. Comme la loi est uniforme, on a

$$E(X) = \sum_{k=a}^{b} k \times \frac{1}{b-a+1} = \frac{1}{b-a+1} \times \left((b-a+1)\frac{a+b}{2} \right) = \frac{a+b}{2}.$$

■

Remarque.
Dans le cas d'une variable suivant une loi uniforme à valeurs dans un intervalle d'entiers, l'espérance est la moyenne des bornes. Mais c'est faux en général. Par exemple, si $X \sim \mathcal{U}(-1, 2, 3)$, alors l'espérance est $-\frac{1}{3} + \frac{2}{3} + \frac{3}{3} = \frac{4}{3} \neq \frac{-1+3}{2}$.

Proposition 3.7 (Espérance d'une loi de Bernoulli)

Soit X une variable aléatoire réelle suivant une loi de Bernoulli de paramètre p. Alors $E(X) = p$.

Démonstration.

L'univers image est ici $\{0, 1\}$, donc $E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = p$. ■

Proposition 3.8 (Espérance d'une loi binomiale)

Soit X une variable aléatoire réelle suivant une loi suivant une loi binomiale de paramètre (n, p). Alors $E(X) = np$.

Démonstration.

Rappelons que l'univers image est $\llbracket 0, n \rrbracket$. Alors

$$E(X) = \sum_{k=0}^{n} kP(X = k) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k}.$$

Or, on a $k\binom{n}{k} = n\binom{n-1}{k-1}$ si $k \in \llbracket 1, n \rrbracket$, et donc

$$\begin{aligned}
E(X) &= \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k(1-p)^{n-k} \\
&= np\sum_{k=1}^{n} \binom{n-1}{k-1}p^{k-1}(1-p)^{n-k} \\
&\overset{\ell=k-1}{=} np\sum_{\ell=0}^{n-1} \binom{n-1}{\ell}p^{\ell}(1-p)^{n-1-\ell} \\
&= np \times \big(p + (1-p)\big)^{n-1} = np
\end{aligned}$$

■

3.3 Linéarité de l'espérance

Proposition 3.9 (Linéarité de l'espérance)

L'espérance est une forme linéaire sur l'espace des variables aléatoires réelles sur Ω, *i.e.* si X et Y sont des variables aléatoires réelles sur Ω, et si $\lambda \in \mathbb{R}$, alors

$$E(X+Y) = E(X) + E(Y), \qquad E(\lambda X) = \lambda E(X).$$

Démonstration.

On applique la proposition 3.2. On a

$$\begin{aligned} E(X+Y) &= \sum_{w\in\Omega} P(\{\omega\})(X+Y)(\omega) = \sum_{w\in\Omega} P(\{\omega\})(X(\omega)+Y(\omega)) \\ &= \sum_{w\in\Omega} P(\{\omega\})X(\omega) + \sum_{w\in\Omega} P(\{\omega\})Y(\omega) = E(X)+E(Y). \end{aligned}$$

On fait de même pour λX. ■

Corollaire 3.10 (Espérance de $aX+b$)

Soient $(a,b)\in\mathbb{R}^2$, et X une variable aléatoire réelle sur Ω. Alors $E(aX+b) = aE(X)+b$.

Démonstration.

On applique la linéarité de l'espérance : $E(aX+b) = aE(X) + E(b) = aE(X)+b$. ■

Corollaire 3.11

Soit X une variable aléatoire réelle sur Ω. Alors $X - E(X)$ est centrée.

Démonstration.

On a $E(X-E(X)) = E(X) - E(X) = 0$ d'après le corollaire précédent. ■

Méthode 3.12 (Décomposition d'une v.a.r. en somme de v.a.r. "simples")

On peut décomposer une v.a.r. X dont on cherche l'espérance en une somme de v.a.r. dont on connait les espérances.

Exemple.

Une urne contient p boules, dont b blanches. On en tire simultanément $n \leqslant p$. On note X le nombre de boules blanches obtenues. On cherche l'espérance de X.

On considère comme univers l'ensemble des parties à n éléments de l'ensemble des boules, muni de la probabilité uniforme. Puis on numérote les boules blanches de 1 à b, et on note X_i la variable indicatrice de l'événement "la boule i est tirée". Son espérance est

$$E(X_i) = P(X_i = 1) = \frac{\binom{p-1}{n-1}}{\binom{p}{n}} = \frac{n}{p},$$

(nombre de cas favorables sur nombre de cas total, puisque'il faut tirer la boule i, accompagnée d'un choix de $n-1$ boules parmi $p-1$). Mais $X = \sum_{i=1}^{b} X_i$, donc

$$E(X) = \sum_{i=1}^{b} E(X_i) = \sum_{i=1}^{b} \frac{n}{p} = \frac{bn}{p}.$$

3.4 Croissance de l'espérance - Inégalité de Markov

Proposition 3.13 (Positivité de l'espérance)

Soit X une variable aléatoire réelle positive sur (Ω, P). Alors :

1. $E(X) \geqslant 0$.
2. $E(X) = 0$ si et seulement si $P(X = 0) = 1$.

Démonstration.

Par définition, si X est positive, son espérance aussi. De plus, une somme de termes positifs est nulle si et seulement si chacun des termes est nul, donc

$$\begin{aligned} E(X) = 0 &\iff \sum_{x \in X(\Omega)} xP(X = x) = 0 \\ &\iff \forall\, x \in X(\Omega),\ xP(X = x) = 0 \\ &\iff \forall\, x \in X(\Omega),\ x \neq 0,\ P(X = x) = 0 \\ &\iff \sum_{x \in X(\Omega)} P(X = x) = P(X = 0) \\ &\iff P(X = 0) = 1, \end{aligned}$$

où la dernière équivalence découle de $\displaystyle\sum_{x \in X(\Omega)} P(X = x) = 1$ (système complet d'événements). ∎

Remarque.

Attention : on peut avoir $P(X = 0) = 1$ sans pour autant que $X = 0$.

Proposition 3.14 (Croissance de l'espérance)

Soient X et Y des variables aléatoires réelles sur (Ω, P), telles que $X \leqslant Y$. Alors $E(X) \leqslant E(Y)$.

Démonstration.

C'est une conséquence de la positivité et de la linéarité : on a $Y - X \geqslant 0$, donc $E(Y - X) \geqslant 0$, *i.e.* $E(Y) - E(X) \geqslant 0$. ∎

Théorème 3.15 (Inégalité de Markov)

Soit X une variable aléatoire réelle **positive** sur (Ω, P). Alors

$$\forall\, a > 0,\ P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

Démonstration.

Montrons que $\mathbb{1}_{(X\geqslant a)} \leqslant \frac{X}{a}$ pour tout $a > 0$. On aura alors (propositions 3.4 et 3.9)

$$P(X \geqslant a) = E\left(\mathbb{1}_{(X\geqslant a)}\right) \leqslant E\left(\frac{X}{a}\right) = \frac{E(X)}{a}.$$

Soit $\omega \in \Omega$. Si $\omega \in (X \geqslant a)$, alors $\mathbb{1}_{(X\geqslant a)}(\omega) = 1 \leqslant \frac{X(\omega)}{a}$. Sinon, comme X est positive, on a $\mathbb{1}_{(X\geqslant a)}(\omega) = 0 \leqslant \frac{X(\omega)}{a}$. ∎

3.5 Formule de transfert

Théorème 3.16 (Formule de transfert)

Soit X une variable aléatoire sur (Ω, P), et $f : X(\Omega) \longrightarrow \mathbb{R}$. Alors

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

ou encore, si $X(\Omega) = \{x_1, \dots, x_n\}$,

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

Démonstration.

D'après la proposition 1.11, on a, pour tout $y \in f(X(\Omega))$,

$$P(f(X) = y) = \sum_{x\in f^{-1}(\{y\})} P(X = x).$$

On en déduit que

$$\begin{aligned} E(f(X)) &= \sum_{y\in f(X)(\Omega)} yP(f(X) = y) = \sum_{y\in f(X)(\Omega)} \left(y \sum_{x\in f^{-1}(\{y\})} P(X = x) \right) \\ &= \sum_{y\in f(X)(\Omega)} \sum_{x\in f^{-1}(\{y\})} yP(X = x) = \sum_{y\in f(X)(\Omega)} \sum_{x\in f^{-1}(\{y\})} f(x)P(X = x) \end{aligned}$$

Mais
$$X(\Omega) = \bigcup_{y \in f(X)(\Omega)} f^{-1}(\{y\}),$$
et cette union est disjointe (un $x \in X(\Omega)$ a une seule image par f), donc
$$\sum_{y \in f(X)(\Omega)} \sum_{x \in f^{-1}(\{y\})} f(x)P(X = x) = \sum_{x \in X(\Omega)} f(x)P(X = x)$$

■

Remarque.

Dans cette formule, ce sont bien les probabilités $P(X = x)$ qui interviennent, et pas $P(f(X) = ..)$. C'est l'intérêt de cette formule : il n'est pas nécessaire de connaître la loi de $f(X)$ pour calculer son espérance.

Exemple.

Soit $X \sim \mathcal{U}(\llbracket -2, 2 \rrbracket)$. Alors $E(X) = 0$, et
$$E(X^2) = \sum_{k=-2}^{2} k^2 P(X = k) = \sum_{k=-2}^{2} k^2 \times \frac{1}{5} = \frac{1}{5} \times (2 \times 2^2 + 2 \times 1^2) = 2.$$

4 Variance, écart type

4.1 Variance - Formule de König-Huygens

Définition 4.1 (Moment d'ordre k)

Soit X une variable aléatoire réelle sur (Ω, P), et $k \in \mathbb{N}$. Le moment d'ordre k de X est $E(X^k)$.

Remarques.

1. Le moment d'ordre 0 est 1, celui d'ordre 1 l'espérance de X.
2. Rappelons que, d'après la formule de transfert,
$$E(X^k) = \sum_{x \in X(\Omega)} x^k P(X = x).$$

Définition 4.2 (Variance)

Soit X une variable aléatoire réelle sur (Ω, P). La variance $V(X)$ de X est le moment d'ordre 2 de $X - E(X)$, *i.e.*
$$V(X) = E\big((X - E(X))^2\big).$$

Remarques.

1. La variance est l'espérance, *i.e.* "la valeur attendue", du carrée de l'écart entre X et son espérance. Elle mesure donc la dispersion des valeurs autour de l'espérance.

2. On prend $(X - E(X))^2$ plutôt que $|X - E(X)|$, car les calculs sont plus simples avec le carré qu'avec la valeur absolue.
3. En appliquant la formule de transfert, on obtient

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

Proposition 4.3 (Formule de König-Huygens)

Soit X une variable aléatoire réelle sur (Ω, P). Alors

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

Démonstration.

On a $(X - E(X))^2 = X^2 - 2XE(X) + E(X)^2$, donc, par linéarité de l'espérance,

$$V(X) = E(X^2) - 2E(X)^2 + E(E(X)^2) = E(X^2) - E(X)^2,$$

car $E(X)^2$ est une v.a.r. certaine, égale à son espérance. ∎

Méthode 4.4

On utilise la formule de König-Huygens pour calculer la variance.

Méthode 4.5

Pour calculer $E(X^2)$, on est parfois amené à calculer $E(X(X+1)$ ou $E(X(X-1))$, ce qui permet d'obtenir $E(X^2)$ par linéarité, en connaissant $E(X)$.

4.2 Propriétés de la variance

Proposition 4.6

Soit X une variable aléatoire réelle sur (Ω, P).

1. $V(X) \geqslant 0$.
2. $V(X) = 0$ si et seulement si $P(X = E(X)) = 1$.

Démonstration.

La variance est l'espérance d'une v.a.r. positive, donc est positive.

Puis, si $P(X = E(X)) = 1$, alors $E(X) \in X(\Omega)$, et par la formule de transfert, $E(X^2) = E(X)^2$, donc $V(X) = 0$.

Réciproquement, si $V(X) = 0$, la variable aléatoire réelle positive $(X - E(X))^2$ a une espérance nulle, donc d'après la proposition 3.13, on a

$$P((X - E(X))^2 = 0) = 1.$$

Or, $(X - E(X))^2 = 0) = (X - E(X) = 0) = (X = E(X))$, donc $P(X = E(X)) = 1$. ∎

Proposition 4.7

Soit X une variable aléatoire réelle sur (Ω, P), et soit $(a, b) \in \mathbb{R}^2$. Alors

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

En particulier, $V(X + b) = V(X)$.

Démonstration.

Par linéarité de l'espérance, on a $E(aX + b) = aE(X) + b$. Un calcul direct donne alors

$$\big((aX + b) - E(aX + b)\big)^2 = a^2\big(X - E(X)\big)^2,$$

et donc

$$V(aX + b) = E\Big(\big((ax + b) - E(aX + b)\big)^2\Big) = E\Big(a^2\big(X - E(X)\big)^2\Big) \overset{\text{lin.}}{=} a^2 E\Big(\big(X - E(X)\big)^2\Big) = a^2 V(X).$$

∎

4.3 Variance des lois usuelles

Proposition 4.8 (Variance d'une variable indicatrice)

Soit A un événement. Alors $V(\mathbb{1}_A) = P(A)(1 - P(A))$.

Démonstration.

On utilise la formule de König-Huygens : $V(X) = E(X^2) - E(X)^2$. Mais ici, $X = X^2$, et $E(X) = P(A)$. ∎

Proposition 4.9 (Variance d'une variable aléatoire certaine)

Soit $a \in \mathbb{R}$. La variance de la variable aléatoire certaine égale à a est 0.

Démonstration.

Déjà vu. ∎

Proposition 4.10 (Variance d'une loi uniforme)

Soit X une variable aléatoire réelle qui suit la loi uniforme sur $[\![a, b]\!]$ ($a, b \in \mathbb{Z}$, $a \leqslant b$). Alors

$$V(X) = \frac{(b - a + 1)^2 - 1}{12}.$$

En particulier, si $X \sim [\![1, n]\!]$ ($n \in \mathbb{N}$), alors $V(X) = \frac{n^2-1}{12}$.

Démonstration.

Traitons d'abord le cas particulier de $a = 1$ et $b = n$. On a $E(X) = \frac{n+1}{2}$, et

$$E(X^2) = \sum_{k=1}^{n} k^2 P(X = k) = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6},$$

et $V(X) = E(X^2) - E(X)^2 = \frac{n^2-1}{12}$.

Pour le cas général, on utilise la proposition 4.7. Si $n = b-a+1$, la variable $X-a+1$ suit la loi uniforme sur $[\![1, n]\!]$, donc

$$V(X-a+1) = V(X) = \frac{n^2-1}{12} = \frac{(b-a+1)^2-1}{12}.$$

■

Remarque.
Il faut se souvenir que $b-a+1$ est le nombre d'éléments de l'intervalle $[\![a, b]\!]$.

Proposition 4.11 (Variance d'une loi de Bernoulli)

Soit X une variable aléatoire réelle suivant une loi de Bernoulli de paramètre p. Alors $V(X) = p(1-p)$.

Démonstration.

L'univers image est ici $\{0, 1\}$, donc $X^2 = X$, et $E(X) = p$, donc

$$V(X) = E(X^2) - E(X)^2 = E(X) - E(X)^2 = p - p^2 = p(1-p).$$

■

Proposition 4.12 (Variance d'une loi binomiale)

Soit X une variable aléatoire réelle suivant une loi suivant une loi binomiale de paramètre (n, p). Alors $V(X) = np(1-p)$.

Démonstration.

Rappelons que l'univers image est $[\![0, n]\!]$, et que $E(X) = np$. On va calculer $E(X^2)$ à partir de $E(X(X-1)$. Rappelons que pour $k \in [\![2, n]\!]$, on a

$$k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}.$$

On a alors, par la formule de transfert,

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} p^k (1-p)^{n-k} \\
&\overset{j=k-2}{=} n(n-1) \sum_{j=0}^{n-2} \binom{n-2}{j} p^{j+2} (1-p)^{n-2-j} \\
&= n(n-1)p^2(p+1-p)^{n-2} = n(n-1)p^2.
\end{aligned}$$

On en déduit que $E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np$, et

$$V(X) = E(X^2) - E(X)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p).$$

■

4.4 Écart type

Définition 4.13 (Écart-type)

Soit X une variable aléatoire réelle sur (Ω, P). Son écart-type est le réel

$$\sigma(X) = \sqrt{V(X)}.$$

Remarques.

1. Grâce à la racine carrée, l'écart-type s'exprime avec les mêmes unités que la variable X.
2. Comme pour la variance, l'écart-type est une mesure de la dispersion autour de l'espérance. Elle mesure aussi la "fiabilité" de l'espérance : si l'écart-type est petit (par rapport aux valeurs prises par X), l'espérance représente bien X.

Définition 4.14 (Variable réduite)

Une variable aléatoire réelle X sur (Ω, P) est réduite si $\sigma(X) = 1$.

Proposition 4.15 (Variable centrée réduite associée à X)

Soit X une variable aléatoire réelle sur (Ω, P) telle que $V(X) > 0$. La variable aléatoire réelle $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite, appelée variable centrée réduite associée à X.

Démonstration.

Par linéarité de l'espérance, on a

$$E\left(\frac{X - E(X)}{\sigma(X)}\right) = \frac{E(X - E(X))}{\sigma(X)} = 0,$$

donc la variable est centrée.

Enfin,

$$V\left(\frac{X - E(X)}{\sigma(X)}\right) = V\left(\frac{X}{\sigma(X)} + b\right) = \frac{V(X)}{\sigma(X)^2} = 1,$$

donc la variable est réduite. ■

4.5 Inégalité de Bienaymé-Tchebychev

Théorème 4.16 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle définie sur (Ω, P). Alors

$$\forall\, \varepsilon > 0,\ P\big(\,|X - E(X)| \geqslant \varepsilon\big) \leqslant \frac{V(X)}{\varepsilon^2}.$$

Démonstration.

La v.a.r. $(X - E(X))^2$ est positive, et son espérance est $V(X)$. D'après l'inégalité de Markov, si $\varepsilon > 0$, on a

$$P((X - E(X))^2 \geqslant \varepsilon^2) \leqslant \frac{V(X)}{\varepsilon^2}.$$

Or, les événements $((X - E(X))^2 \geqslant \varepsilon^2)$ et $(|X - E(X)| \geqslant \varepsilon)$ sont égaux, d'où l'inégalité. ■

Remarques.

1. On retourve que la variance mesure la dispersion de X autour de son espérance. Ainsi, pour tout $\varepsilon > 0$, la probabilité que l'écart de X à $E(X)$ soit supérieur à ε est d'autant plus petite que la variance de X sera faible.
2. L'événement $(|X - E(X)| < \varepsilon)$ est l'événement contraire de $(|X - E(X)| \geqslant \varepsilon)$. L'inégalité peut donc se réécrire

$$\forall\, \varepsilon > 0,\ P\big(\,|X - E(X)| < \varepsilon\big) \geqslant 1 - \frac{V(X)}{\varepsilon^2}.$$

L'inégalité montre donc que X va prendre des valeurs proches de son espérance avec une grande probabilité.

5 Couples de variables aléatoires

5.1 Définitions

Définition 5.1 (Couple de variables aléatoires)

Soient $X : \Omega \longrightarrow E$ et $Y : \Omega \longrightarrow E'$ des variables aléatoires. Le couple des variables aléatoires (X, Y) est la fonction

$$\begin{array}{rcl} \Omega & \longrightarrow & E \times E' \\ \omega & \longmapsto & (X(\omega), Y(\omega)). \end{array}$$

Lorsque $E = E' = \mathbb{R}$, on a un couple de variables aléatoires réelles.

Exemple.

On tire d'une urne contenant 4 boules blanches, 5 noires et 6 rouges 5 boules. On note X le nombre de boules blanches, Y le nombre de boules noires. Alors (X, Y) est un couple de v.a.r.

Remarque.

Attention, l'univers image n'est pas $X(\Omega) \times Y(\Omega)$ en général, mais seulement $\{(X(\omega), Y(\omega),\ \omega \in \Omega\}$.

Voici un exemple : on lance un dé. On note X le reste modulo 2 du chiffre obtenu, et Y le reste modulo 4. Alors

$$\Omega = [\![1, 6]\!], \qquad X(\Omega) = \{0, 1\}\,, \qquad Y(\Omega) = \{0, 1, 2, 3\}\,,$$

et

$$X(\Omega) \times Y(\Omega) = \{(0, 0); (0, 1); (0, 2); (0, 3); (1, 0); (1, 1); (1; 2); (1, 3)\}\,.$$

On a aussi

$$(X, Y)(\Omega) = \{(0, 0); (0, 2); (1, 1); (1, 3)\}\,,$$

comme on peut le vérifier en calculant $(X(i), Y(i))$ pour $i \in [\![1, 6]\!]$.

Proposition 5.2 (Système complet d'événements d'un couple de v.a.r.)

Soit (X, Y) un couple de variables aléatoires réelles sur Ω. La famille $\Big((X = x) \cap (Y = y)\Big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'événements, appelé système complet d'événement associé au couple (X, Y).

Démonstration.

Les ensembles $(X = x) \cap (Y = y)$ sont bien des événements. Ils sont deux à deux incompatibles puisque si $(x, y) \neq (x', y')$, alors par exemple $x \neq x'$, et donc $(X = x) \cap (Y = y) \cap (X = x') \cap (Y = y') = \emptyset$. Enfin, si $\omega \in \Omega$, alors $\omega \in (X = X(\omega)) \cap (Y = Y(\omega))$, donc

$$\Omega = \bigcup_{(x,y)\in X(\Omega)\times Y(\Omega)} \,.$$

■

Remarque.

Si $X(\Omega) = \{x_1, \ldots, x_n\}$ et $Y(\Omega) = \{y_1, \ldots, y_m\}$, La famille s'écrit alors $((X = x_i) \cap (Y = y_j))_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$.

Corollaire 5.3

Soit (X, Y) un couple de variables aléatoires réelles sur Ω. Alors

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} P((X = x) \cap (Y = y)) = 1,$$

ou encore, si $X(\Omega) = \{x_1, \ldots, x_n\}$ et $Y(\Omega) = \{y_1, \ldots, y_m\}$,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P((X = x_i) \cap (Y = y_j)) = 1.$$

Démonstration.

C'est une conséquence de la proposition 5.2 : la somme des probabilités d'un système complet d'événements vaut 1. ■

Remarques.

1. On note en général $(X = x) \cap (Y = y) = (X = x, Y = y)$, (remarquez qu'on a aussi $(X = x, Y = y) = ((X, Y) = (x, y))$, et de même $P((X = x) \cap (Y = y)) = P(X = x, Y = y)$.
2. Bien entendu, si $(x, y) \notin (X, Y)(\Omega)$, alors $P(X = x, Y = y) = 0$.

5.2 Loi conjointe et lois marginales

Définition 5.4 (Loi conjointe)

Soient X et Y deux variables aléatoires sur l'espace probabilisé (Ω, P). La loi conjointe de X et Y est la probabilité sur $X(\Omega) \times Y(\Omega)$ définie par

$$\begin{array}{ccc} X(\Omega) \times Y(\Omega) & \longrightarrow & [0, 1] \\ (x, y) & \longmapsto & P(X = x, Y = y). \end{array}$$

Démonstration.

Il faut bien entendu montrer qu'une telle formule définie bien une probabilité sur Ω. Pour cela, on applique le théorème 2.7 du chapitre 26 : on définit les probabilités des événements élémentaires de $X(\Omega) \times Y(\Omega)$ avec les réels positifs $P(X = x, Y = y)$ dont la somme vaut 1 (corollaire 5.3). ■

Remarques.

1. Si on note $X(\Omega) = \{x_1, \dots, x_p\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$, la loi est déterminée par la famille

$$(P(X = x_i, Y = y_j))_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant m}} .$$

On se donne la probabilité de tous les couples possibles.

2. On note souvent la loi conjointe sous forme d'un tableau à double entrée, *cf.* les exemples.
3. Notons $Z(\Omega)$ l'univers image du couple (X, Y). On a vu dans un exemple plus haut qu'en général, $Z(\Omega) \neq X(\Omega) \times Y(\Omega)$. En particulier, si $(x, y) \in X(\Omega) \times Y(\Omega) \setminus Z(\Omega)$, alors $P(X = x, Y = y) = 0$.

Exemples.

1. Reprenons l'exemple de lancé du dé, avec X le reste modulo 2, et Y celui modulo 4, et équiprobabilité de chaque numéro. Alors

$$X(\Omega) \times Y(\Omega) = \{(0, 0); (0, 1); (0, 2); (0, 3); (1, 0); (1, 1); (1; 2); (1, 3)\} .$$

On vérifie que

$$P(X = 0, Y = 1) = P(X = 0, Y = 3) = P(X = 1, Y = 0) = P(X = 1, Y = 2) = 0,$$

puis que

$$P(X = 0, Y = 0) = P(X = 1, Y = 3) = \frac{1}{6}, \qquad P(X = 0, Y = 2) = P(X = 1, Y = 1) = \frac{1}{3}.$$

En effet, on a par exemple $(X = 0, Y = 2) = \{2; 6\}$. Sous forme de tableau, on obtient

$X \backslash Y$	0	1	2	3
0	$\frac{1}{6}$	0	$\frac{1}{3}$	0
1	0	$\frac{1}{3}$	0	$\frac{1}{6}$

2. Calculons la loi de la somme $X+Y$. On a $(X+Y)(\Omega) = [\![0,4]\!]$. Pour $k \in [\![0,4]\!]$, notons

$$S_k = \{(i,j) \in X(\Omega) \times Y(\Omega),\ i+j=k\} = \{(i,j) \in [\![0,1]\!] \times [\![0,3]\!],\ i+j=k\}.$$

Alors

$$(X+Y=k) = \bigcup_{(i,j)\in S_k} (X=i, Y=j),$$

et comme ces événements sont deux à deux incompatibles, on a

$$P(X+Y=k) = \sum_{(i,j)\in S_k} P(X=i, Y=j).$$

Par exemple,

$$P(X+Y=2) = P(X=0, Y=2) + P(X=1, Y=1) = \frac{2}{3}.$$

Définition 5.5 (Lois marginales)

Soit (X,Y) un couple de variables aléatoires. La loi de X est la première loi marginale de (X,Y), et la loi de Y est la deuxième loi marginale de (X,Y).

Théorème 5.6 (Lois marginales à partir de la loi conjointe)

Soit (X,Y) un couple de variables aléatoires sur Ω. Si $X(\Omega) = \{x_1, \dots, x_p\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$, alors

$$\forall\, i \in [\![1,p]\!],\ P(X=x_i) = \sum_{j=1}^{m} P(X=x_i, Y=y_j),$$
$$\forall\, j \in [\![1,m]\!],\ P(Y=y_j) = \sum_{i=1}^{p} P(X=x_i, Y=y_j).$$

Démonstration.

Prouvons la première égalité. Elle résulte du fait que la famille $((Y=y_j))_{1\leqslant j\leqslant m}$ est un système complet d'événements. On a donc, pour $i \in [\![1,p]\!]$ fixé,

$$P(X=x_i) = \sum_{j=1}^{m} P((X=x_i)\cap(Y=y_j)) = \sum_{j=1}^{m} P(X=x_i, Y=y_j).$$

■

Remarques.

1. On peut aussi écrire le théorème sous la forme suivante :

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

2. Les lois marginales ne suffisent pas, en général, à reconstituer la loi conjointe du couple, *cf.* les exemples.
3. Si on écrit la loi conjointe dans un tableau, les lois marginales s'obtiennent en sommant les lignes ou les colonnes.

Exemple.

Reprenons l'exemple de lancé du dé, avec X le reste modulo 2, et Y celui modulo 4, et équiprobabilité de chaque numéro. Le tableau donne alors

$X\backslash Y$	0	1	2	3	loi de X
0	$\frac{1}{6}$	0	$\frac{1}{3}$	0	$\frac{1}{2}$
1	0	$\frac{1}{3}$	0	$\frac{1}{6}$	$\frac{1}{2}$
loi de Y	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	1

On vérifie que la somme de la dernière ligne, et de la dernière colonne, valent bien 1.

Remarque.

Exemple où la connaissance des lois marginales ne suffit pas pour retrouver la loi du couple.

À un oral de concours, 3 sujets sont possibles : A, B ou C. Deux candidats participent à ce concours et tout deux ne maîtrisent que le sujet A. Notons X la v.a. modélisant le succès du premier candidat ($X = 1$ si succès, 0 si échec) et Y celle modélisant le succès du second candidat.

On suppose dans un premier temps que le sujet est remis dans le panier des sujets après chaque tirage. Le tableau donnant la loi conjointe est le suivant :

$X\backslash Y$	0	1	loi de X
0	4/9	2/9	2/3
1	2/9	1/9	1/3
loi de Y	2/3	1/3	1

On suppose maintenant que le tirage du sujet se fait sans remise. En utilisant les probabilités composées, on a :

$$\mathbb{P}(X = 0 \cap Y = 0) = \mathbb{P}(X = 0)\mathbb{P}_{X=0}(Y = 0) = 2/3 \times 1/2 = 1/3,$$
$$\mathbb{P}(X = 1 \cap Y = 0) = 1/3 \times 1 = 1/3,$$
$$\mathbb{P}(X = 1 \cap Y = 1) = 0 \text{ et } \mathbb{P}(X = 0 \cap Y = 1) = 1/3.$$

La loi conjointe et les loi marginales sont donc données par le tableau suivant :

$X\backslash Y$	0	1	loi de X
0	1/3	1/3	2/3
1	1/3	0	1/3
loi de Y	2/3	1/3	1

On trouve les mêmes lois marginales, mais la loi du couple est différente.

5.3 Espérance d'un produit

Proposition 5.7 (Espérance d'un produit)

Soient X et Y des variables aléatoires réelles sur (Ω, P). Alors

$$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x, Y = y),$$

ou encore, si $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$,

$$E(XY) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P(X = x_i, Y = y_j).$$

Démonstration.

Soit $Z = (X, Y)$. Alors $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$, et $XY = f(Z)$, où f est la fonction $(x, y) \longmapsto xy$. En appliquant la formule de transfert à $f(Z)$, on obtient

$$E(XY) = \sum_{z \in Z(\Omega)} f(z)P(Z = z) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i, y_j)P((X, Y) = (x_i, y_j)),$$

car $P((X, Y) = (x_i, y_j)) = 0$ si $(x_i, y_j) \in X(\Omega) \times Y(\Omega) \setminus Z(\Omega)$. Cela donne la formule voulue. ■

5.4 Lois conditionnelles

Définition 5.8 (Lois conditionnelles)

Soit (X, Y) un couple de variables aléatoires sur l'espace probabilisé (Ω, P).

1. Soit $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$. La loi de X sachant $(Y = y)$ est la loi de X dans l'espace probabilisé $(\Omega, P_{(Y=y)})$. Elle est donc déterminée pour tout $x \in X$ par :

$$P_{(Y=y)}(X = x) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

2. Soit $x \in X(\Omega)$ tel que $P(X = x) \neq 0$. La loi de Y sachant $(X = x)$ est la loi de Y dans l'espace probabilisé $(\Omega, P_{(X=x)})$. Elle est donc déterminée pour tout $y \in Y$ par :

$$P_{(X=x)}(Y = y) = P(Y = y|X = x) = \frac{P(Y = y, X = x)}{P(X = x)}.$$

Remarques.

1. Rappelons que la probabilité $P_{(Y=y)}$ est la probabilité conditionnelle à l'événement $(Y = y)$.
2. On parle aussi de loi conditionelle à $(Y = y)$.

Exemple.

On lance deux dés à 4 faces de manière indépendante (l'univers est donc $\Omega = [\![1, 4]\!]^2$). On note X le résultat du premier lancer, Y le second. On pose $M = \max(X, Y)$ et on s'intéresse à la loi du couple $Z = (X, M)$. On a $Z(\Omega) \subset [\![1; 4]\!]^2$ (on va voir qu'il n'y a pas égalité).

Soit $(i, j) \in Z(\Omega)$.

Si $i > j$, $(X = i) \cap (M = j) = (X = i) \cap (\max(X, Y) = j) = \emptyset$,donc $\mathbb{P}(X = i, M = j) = 0$.

—Si $i = j$, $(X = i) \cap (M = j) = (X = i) \cap (Y \leqslant i)$, donc $\mathbb{P}(X = i) \cap (M = i) = \mathbb{P}(X = i)\mathbb{P}(Y \leqslant i) = \frac{1}{4} \times \frac{i}{4}$.

—Si $i < j$, $(X = i) \cap (M = j) = (X = i) \cap (Y = j)$, d'où $\mathbb{P}(X = i) \cap (M = j) = \frac{1}{4} \times \frac{1}{4}$.

La loi conjointe et les lois marginales sont donc :

$X\backslash M$	1	2	3	4	loi de X
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	1/4
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	1/4
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$	1/4
4	0	0	0	$\frac{4}{16}$	1/4
loi de M	1/16	3/16	5/16	7/16	

La loi de X sachant $(M = 2)$ s'obtient en regardant la colonne $(M = 2)$, et à diviser par la probabilité $P(M = 2)$ qui est dans la ligne "loi de M". Par exemple,

$$P_{(M=2)}(X = 1) = \frac{1/16}{3/16} = \frac{1}{3}.$$

Théorème 5.9 (Loi conjointe par loi marginale et lois conditionnelles)

Soit (X, Y) un couple de variables aléatoires sur Ω.

1. $P(X = x) = \displaystyle\sum_{y \in Y(\Omega)} P(X = x, Y = y).$

2. $P(Y = y) = \displaystyle\sum_{x \in X(\Omega)} P(X = x, Y = y).$

On a de plus :

3. $P(X = x, Y = y) = \begin{cases} P(Y = y)P(X = x|Y = y) & \text{si } P(Y = y) \neq 0 \\ P(X = x)P(Y = y|X = x) & \text{si } P(X = x) \neq 0 \end{cases}$

Et enfin :

4. $P(X = x) = \displaystyle\sum_{y \in Y(\Omega)} P(Y = y)P(X = x|Y = y)$ si pour tout $y \in Y(\Omega)$, $P(Y = y) \neq 0$.

5. $P(Y = y) = \displaystyle\sum_{x \in X(\Omega)} P(X = x)P(Y = y|X = x)$ si pour tout $x \in X(\Omega)$, $P(X = x) \neq 0$.

Démonstration.

Le 1 découle de la définition des probabilités conditionnelles. Les points 2 et 3 découlent directement du théorème 5.6 et du point 1. ■

Remarque.

On en déduit que la loi marginale d'une des variable associée à sa probabilité conditionnelle donne la loi conjointe.

6 Variables aléatoires indépendantes

Définition 6.1 (Couple de variables aléatoires indépendantes)

Soit (X, Y) un couple de variables aléatoires sur (Ω, P). Elles sont indépendantes pour la probabilité P si pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants pour la probabilité P, *i.e.* si

$$\forall\,(A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)),\ P(X \in\ A, Y \in\ B) = P(X \in\ A)P(Y \in\ B).$$

Proposition 6.2

Soit (X, Y) un couple de variables aléatoires sur (Ω, P). Elles sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants, *i.e.*

$$\forall\,(x, y) \in X(\Omega) \times Y(\Omega),\ P(X = x, Y = y) = P(X = x)P(Y = y).$$

Démonstration.

Si X et Y sont indépendantes, les événements $A = \{x\}$ et $B = \{y\}$ sont bien indépendants.

Réciproquement, supposons que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants. Soient $A \subset X(\Omega)$ et $B \subset Y(\Omega)$. On a donc

$$(X \in A) \cap (Y \in B) = \bigcup_{x \in A} \bigcup_{y \in B} (X = x) \cap (Y = y).$$

Mais tous ces événements sont incompatibles, et de plus indépendants par hypothèse, et en nombre fini, donc

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A} \sum_{y \in B} P(X = x, Y = y) \\ &= \sum_{x \in A} \sum_{y \in B} P(X = x)P(Y = y) \\ &= \left(\sum_{x \in A} P(X = x)\right)\left(\sum_{y \in B} P(Y = y)\right) \\ &= P(X \in A)P(Y \in B). \end{aligned}$$

Les événements sont bien indépendants. ■

Remarques.

1. Si (X, Y) sont indépendantes, les lois marginales déterminent entièrement la loi conjointe (par produit).
2. Comme pour les événements, l'indépendance dépend de la probabilité choisie. Prenons par exemple $\Omega = \{1, 2\}^2$. On note X la première coordonnée, Y la deuxième. On note P la probabilité uniforme sur Ω, et P' la probabilité définie par

$$P'((i, j) = \begin{cases} \frac{1}{3} & \text{si } i = j, \\ \frac{1}{6} & \text{sinon.} \end{cases}$$

C'est une probabilité car la somme des probabilités des événements élémentaires vaut 1. On a alors

$$P(X = i, Y = j) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(X = i)P(Y = j).$$

Mais
$$P'(X=1,Y=1)=\frac{1}{3}\neq P'(X=1)P'(Y=1)=\frac{1}{2}\times\frac{1}{2},$$
par la formule des probabilités totales (ou : mêmes lois marginales que P).

Proposition 6.3

Soit (X,Y) un couple de variables aléatoires sur (Ω,P). Elles sont indépendantes si et seulement si au moins une des deux conditions suivantes est vérifiée :

1. $\forall\,(x,y)\in X(\Omega)\times Y(\Omega),\ P(Y=y)\neq 0\Longrightarrow P(X=x|Y=y)=P(X=x)$.
2. $\forall\,(x,y)\in X(\Omega)\times Y(\Omega),\ P(X=x)\neq 0\Longrightarrow P(Y=y|X=x)=P(Y=y)$.

Démonstration.

Montrons l'équivalence avec le point 1.

Si X et Y sont indépendantes, et $(x,y)\in X(\Omega)\times Y(\Omega)$, et $P(Y=y)\neq 0$, alors
$$P(X=x|Y=y)=\frac{P(X=x,Y=y)}{P(Y=y)}\overset{\text{ind.}}{=}\frac{P(X=x)P(Y=y)}{P(Y=y)}=P(X=x).$$
(On peut aussi utiliser la proposition 4.2 du chapitre 26 appliquée aux événements $A=(Y=y)$ et $B=(X=x)$).

—Traitons la réciproque. On suppose que 1 est vérifié. Soit $(x,y)\in X(\Omega)\times Y(\Omega)$. Montrons alors que $P(X=x,Y=y)=P(X=x)P(Y=y)$. Cela prouvera que X et Y sont indépendantes (proposition 6.2).

Si $P(Y=y)=0$, on a $P(X=x,Y=y)=0$ (car $(X=x,Y=y)\subset(Y=y)$), donc $P(X=x,Y=y)=P(X=x)P(Y=y)$. Sinon, on raisonne comme pour la proposition 4.2 du chapitre 26 appliquée aux événements $A=(Y=y)$ et $B=(X=x)$.

La proposition est démontrée. ∎

Proposition 6.4 (Images de variables indépendantes)

Soient X et Y des variables aléatoires indépendantes sur (Ω,P). Soit f une fonction définie sur $X(\Omega)$ et g une fonction définie sur $Y(\Omega)$. Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes pour la probabilité P.

Démonstration.

Prenons par exemple des fonctions f et g à valeurs réelles. Soient A et B des parties de $\mathbb{R}$. Alors
$$\begin{aligned}P(f(X)\in A,g(Y)\in B)&=P(X\in f^{-1}(A),Y\in g^{-1}(B))\\&=P(X\in f^{-1}(A))P(Y\in g^{-1}(B))\\&=P(f(X)\in A)P(g(Y)\in B).\end{aligned}$$

Les variables $f(X)$ et $g(Y)$ sont bien indépendantes. ∎

Remarque.

Traitons le cas de la somme de deux variables indépendantes. On a

$$P(X+Y=z) = \sum_{x+y=z} P(X=x, Y=y) = \sum_{x+y=z} P(X=x)P(Y=y).$$

Dans le cas où X et Y sont à valeurs dans $\mathbb{N}$, cela donne

$$P(X+Y=k) = \sum_{i=0}^{k} P(X=i)P(Y=k-i).$$

Proposition 6.5 (Espérance d'un produit de variables indépendantes)

Soient X et Y des variables aléatoires réelles indépendantes sur (Ω, P). Alors $E(XY) = E(X)E(Y)$.

Démonstration.

Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. Comme X et Y sont indépendantes, on a

$$P(X=x, Y=y) = P(X=x)P(Y=y),$$

et donc, si $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$, on a

$$\begin{aligned}
E(XY) &= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P(X=x_i, Y=y_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P(X=x_i)P(Y=y_j) \\
&= \left(\sum_{i=1}^{n} x_i P(X=x_i) \right) \left(\sum_{j=1}^{m} y_j P(Y=y_j) \right) \\
&= E(X)E(Y).
\end{aligned}$$

■

Remarque.

Attention : la réciproque est fausse : si $E(XY) = E(X)E(Y)$, en général, les variables X et Y ne sont pas indépendantes. Si $X \sim \mathcal{U}([\![-1, 1]\!])$ et $Y = X^2$, on a $E(X) = 0$, $E(Y) = E(X^2) = \frac{2}{3}$ (utilisez la formule de transfert), et $E(XY) = E(X^3) = 0$ (de même, formule de transfert), donc $E(XY) = E(X)E(Y)$. Pourtant, $P(X=1, Y=0) = 0 \neq P(X=1)P(X^2=0)$.

7 Covariance

7.1 Définition

Définition 7.1 (Covariance de deux variables aléatoires réelles)

Soient X et Y deux variables aléatoires réelles définies sur un même espace probabilisé (Ω, P). La covariance de X et Y est le réel

$$\mathrm{Cov}(X, Y) = E\Big((X - E(X))(Y - E(Y))\Big).$$

Remarques.

1. C'est l'espérance du produit des variables aléatoires centrées associées à X et Y.
2. La covariance mesure combien les variables sont corrélées. Si elle est > 0, les variables varient dans le même sens (si l'une est grande, l'autre aussi).
3. On a bien sûr $\mathrm{Cov}(X, X) = V(X)$.

Proposition 7.2

Soient X et Y deux variables aléatoires réelles définies sur un même espace probabilisé (Ω, P). Alors

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Démonstration.

Par linéarité, on a

$$\begin{aligned}\mathrm{Cov}(X, Y) = E(XY - E(Y)X - E(X)Y + E(X)E(Y)) = E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y)\\ = E(XY) - E(X)E(Y).\end{aligned}$$

■

Remarque.

On rappelle qu'on a déjà vu comment calculer l'espérance d'un produit, *cf.* la proposition 5.7.

Exemple.

On reprend un exemple vu au paragraphe 5.4. On lance deux dés à 4 faces de manière indépendante. On note X le résultat du premier lancer, Y le second. On pose $M = \max(X, Y)$. La loi conjointe et les lois marginales sont données par

$X\backslash M$	1	2	3	4	loi de X
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	1/4
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	1/4
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$	1/4
4	0	0	0	$\frac{4}{16}$	1/4
loi de M	1/16	3/16	5/16	7/16	

Déterminons maintenant $E(XM)$ grâce à la proposition 5.7 : c'est une somme à $4 \times 4 = 16$ termes (puisque $Z(\Omega) = [\![1, 4]\!]^2$), mais il y a 6 probabilités nulles, donc il reste une somme à 10 termes, et on obtient $E(XM) = \frac{135}{16}$. Les lois marginales nous donnent aussi $E(X) = \frac{5}{2}$ et $E(M) = \frac{5}{2}$, donc $\mathrm{Cov}(X, M) = \frac{35}{16} > 0$: M a d'autant plus de chance d'être grand si X l'est.

7.2 Propriétés de la covariance

Proposition 7.3 (Bilinéarité)

La covariance est une forme bilinéaire symétrique positive sur l'espace des variables aléatoires réelles sur Ω.

Démonstration.

C'est une conséquence de la linéarité de l'espérance. ■

Remarque.
Attention : elle n'est pas définie positive, car $V(X) = 0$ n'implique pas $X = 0$, mais seulement $P(X = 0) = 1$.

Proposition 7.4 (Cas de variables indépendantes)

Soient X et Y deux variables aléatoires réelles indépendantes définies sur un même espace probabilisé (Ω, P). Alors $\mathrm{Cov}(X, Y) = 0$.

Démonstration.

On utilise les propositions 6.5 et 7.2. ■

Remarques.

1. La réciproque est fausse. Il suffit de reprendre l'exemple de deux variables aléatoires X et Y non indépendantes telles que $E(XY) = E(X)E(Y)$, juste après la proposition 6.5.
2. Deux variables dont la covaraince est nulle sont dîtes non corrélées.

7.3 Variance d'une somme

Théorème 7.5 (Variance d'une somme de deux variables)

Soient X et Y deux variables aléatoires réelles définies sur un même espace probabilisé (Ω, P). Alors

$$V(X + Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y).$$

Démonstration.

On utilise la bilinéarité et la symétrie de la covariance : on a

$$V(X+Y) = \mathrm{Cov}(X+Y, X+Y) = \mathrm{Cov}(X, X)+\mathrm{Cov}(Y, Y)+\mathrm{Cov}(X, Y)+\mathrm{Cov}(Y, X) = V(X)+V(Y)+2\mathrm{Cov}(X, Y).$$

■

Remarque.
On en déduit que $V(aX + bY) = a^2V(X) + b^2V(Y) + 2ab\mathrm{Cov}(X, Y)$.

Théorème 7.6 (Variance d'une somme de deux variables indépendantes)

Soient X et Y deux variables aléatoires réelles indépendantes définies sur un même espace probabilisé (Ω, P). Alors

$$V(X + Y) = V(X) + V(Y).$$

Démonstration.

D'après la proposition 7.4 et le théorème 7.5, si X et Y sont indépendantes, $\mathrm{Cov}(X, Y) = 0$, donc $V(X+Y) = V(X) + V(Y)$. ■

8 Extension aux n-uplets de variables aléatoires

8.1 Définitions

Définition 8.1 (n-uplet de variables aléatoires)

Soit $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires sur Ω à valeurs respectivement dans $E_1, \ldots, E_n$. Le n-uplet de variables aléatoires $(X_1, \ldots, X_n)$ est la variable aléatoire

$$\begin{array}{ccc} \Omega & \longrightarrow & E_1 \times \cdots \times E_n \\ \omega & \longmapsto & \Big(X_1(\omega), \ldots, X_n(\omega)\Big). \end{array}$$

Si $E_k = \mathbb{R}$ pour tout k, $(X_1, \ldots, X_n)$ est un n-uplet de variables aléatoires réelles.

Proposition 8.2

Soit $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires sur Ω. La famille

$$\Big((X_1 = x_1) \cap \cdots \cap (X_n = x_n)\Big)_{(x_1,\ldots,x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)}$$

forme un système complet d'événements de Ω.

Démonstration.

La démonstration est la même que celle de la proposition 5.2. ■

Définition 8.3 (Loi conjointe, lois marginales)

Soit $(X_1, \ldots, X_n)$ un n-uplet de variables aléatoires sur l'espace probabilisé (Ω, P). La loi conjointe de $X_1, \ldots, X_n$ est la probabilité sur $X_1(\Omega) \times \cdots \times X_n(\Omega)$ définie par

$$\begin{array}{ccc} X_1(\Omega) \times \cdots \times X_n(\Omega) & \longrightarrow & [0, 1] \\ (x_1, \ldots, x_n) & \longmapsto & P(X_1 = x_1, \ldots, X_n = x_n). \end{array}$$

Les lois marginales sont les lois de $X_1, \ldots, X_n$.

Remarque.

C'est bien une probabilité car, comme pour les couples de variables aléatoires, elle est définie sur les événements élémentaires de $X_1(\Omega) \times \cdots \times X_n(\Omega)$, et la somme des probabilités vaut 1 (système complet d'événements).

Théorème 8.4 (Lois marginales)

Soit $(X_1, \dots, X_n)$ un n-uplet de variables aléatoires sur l'espace probabilisé (Ω, P). Pour $i \in [\![1, n]\!]$, notons $n_i = \text{card}(X_i(\Omega))$, et $X_i(\Omega) = \{x_{i1}, \dots, x_{in_i}\}$. On a alors pour tout $j \in [\![1, n_i]\!]$,

$$P(X_i = x_{ij}) = \sum_{\substack{1 \leqslant j_k \leqslant n_k \\ k=1,\dots,n,\ k \neq i}} P(X_1 = x_{1j_1}, \dots, X_{i-1} = x_{i-1,j_{i-1}}, X_i = x_{ij}, X_{i+1} = x_{i+1,j_i+1}, \dots, X_n = x_{nj_n}).$$

Démonstration.

C'est encore une fois la formule des probabilités totales. Faîtes la démonstration si tout le reste est acquis. ■

Remarque.

La somme se fait en fixant $X_i = x_{ij}$.

Exemple.

On considère une urne contenant p boules numérotées de 1 à p. On en tire simultanément n boules ($n \leqslant p$). Les numéros obtenus sont notés $Y_1 < Y_2 < \cdots < Y_n$. On veut la loi du n-uplet $Z = (Y_1, \dots, Y_n)$, et les lois marginales.

Soit $(y_1, \dots, y_n) \in [\![1, p]\!]^n$. Par définition des Y_k, si deux des y_i sont égaux, ou si le n-uplet n'est pas rangé dans l'ordre croissant, on a $P(Y_1 = y_1, \dots, Y_n = y_n) = 0$. Sinon, il y a $\binom{p}{n}$ tirages possibles, équiprobables, chacun donnant un et un seul n-uplet rangé dans l'ordre strictement croissant, donc

$$y_1 < \cdots < y_n \Longrightarrow P(Y_1 = y_1, \dots, Y_n = y_n) = \frac{1}{\binom{p}{n}}.$$

On a la loi conjointe de Z.

Déterminons maintenant les lois marginales. Fixons $k \in [\![1, n]\!]$. L'ensemble des valeurs prises par Y_k est $[\![k, p-n+k]\!]$ (de $(Y_1, \dots, Y_k) = (1, \dots, k)$ à $(Y_k, \dots, Y_n) = (p-n+k, \dots, p)$). Fixons alors $y_k \in [\![k, p-n+k]\!]$. Un n-uplet $(y_1, \dots, y_n)$ a une probabilité non nulle de sortir si et seulement si $1 \leqslant y_1 < \cdots y_{k-1} < y_k < y_{k+1} < \cdots < y_n$. Il y a donc $\binom{y_k-1}{k-1}$ choix possible pour $(y_1, \dots, y_{k-1})$, et $\binom{p-y_k}{n-k}$ choix possibles pour $(y_{k+1}, \dots, y_n)$, donc finalement $\binom{y_k-1}{k-1}\binom{p-y_k}{n-k}$ n-uplets à probabilité non nulle, et toutes égales, donc

$$P(Y_k = y_k) = \frac{\binom{y_k-1}{k-1}\binom{p-y_k}{n-k}}{\binom{p}{n}}.$$

8.2 Indépendance de n variables aléatoires

Définition 8.5 (Indépendance mutuelle de variables aléatoires)

Des variables aléatoires $X_1, \dots, X_n$ sur (Ω, P) sont mutuellement indépendantes si pour tout $A_i \subset X_i(\Omega)$ $(i = 1, \dots, n)$, les événements $(X_1 \in A_1), \dots, (X_n \in A_n)$ sont indépendants, *i.e.* si

$$\forall\, A_1 \times \cdots \times A_n \in \mathcal{P}(X_1(\Omega)) \times \cdots \times \mathcal{P}(X_n(\Omega)),\ P(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i).$$

Remarques.

1. On dit aussi simplement "indépendantes".
2. Rappelons que la notation $A_1 \times \cdots \times A_n \in \mathcal{P}(X_1(\Omega)) \times \cdots \times \mathcal{P}(X_n(\Omega))$ signifie que pour tout $i = 1, \ldots, n$, $A_i \subset X_i(\Omega)$.

Proposition 8.6

Des variables aléatoires $X_1, \ldots, X_n$ sur (Ω, P) sont mutuellement indépendantes si, et seulement si, pour tout $(x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, les événements $(X_1 = x_1), \ldots, (X_n = x_n)$ sont indépendants, *i.e.* si, et seulement si,

$$\forall\, (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega),\ P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i).$$

Démonstration.

Comme pour les couples. Faîtes-là si vous avez le temps. Sinon, admis. ■

Théorème 8.7 (Variance d'une somme)

Soient $X_1, \ldots, X_n$ des variables aléatoires réelles sur (Ω, P). Alors

$$V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \operatorname{Cov}(X_i, X_j).$$

Démonstration.

C'est une récurrence sur n. Faîtes le si vous avez le temps. ■

Théorème 8.8 (Variance d'une somme de variables deux à deux indépendantes)

Soient $X_1, \ldots, X_n$ des variables aléatoires réelles deux à deux indépendantes sur (Ω, P). Alors

$$V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i).$$

Démonstration.

C'est le même raisonnement. On utilise la proposition 7.4 et le théorème 8.7. ■

Proposition 8.9 (Sommes de variables de Bernoulli indépendantes)

Soient $X_1, \ldots, X_n$ des variables aléatoires réelles de Bernoulli, de paramètre p, deux à deux indépendantes sur (Ω, P). Alors

$$V(X_1 + \cdots + X_n) = np(1 - p).$$

Démonstration.

Chacune a une variance de $p(1-p)$. Ou encore $X_1 + \cdots + X_n \sim B(n,p)$. ■

Remarque.

Si on sait décomposer une variable binomiale de paramètre (n,p) en somme de variables de Bernoulli deux à deux indépendantes, on trouve facilement la variance de la variable binomiale.