

Chapitre 27

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} .

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 .

Définition 1.3 (Variable indicatrice d'un événement)

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

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

Remarques.

1. L'ensemble des variables aléatoires réelles est l'ensemble \mathbb{R}^Ω , 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 $f(X)$.

Remarque.

La notation $u(X)$ est cohérente avec le vocabulaire "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
 - $(X = x) = \{X = x\} = \{\omega \in \Omega, X(\omega) = x\} = X^{-1}(\{x\})$.
 - $(X \leq x) = \{X \leq x\} = \{\omega \in \Omega, X(\omega) \leq x\} = X^{-1}(]-\infty, x])$.
 - $(X \geq x) = \{X \geq x\} = \{\omega \in \Omega, X(\omega) \geq x\} = X^{-1}([x, +\infty[)$.
 - $(X < x) = \{X < x\} = \{\omega \in \Omega, X(\omega) < x\} = X^{-1}(]-\infty, x[)$.
 - $(X > x) = \{X > x\} = \{\omega \in \Omega, X(\omega) > x\} = X^{-1}(]x, +\infty[)$.
 (Ce sont des événements.)
3. On note $P(X \in A) = P(\{X \in A\})$ et $P(X = x) = P(\{X = x\})$.

Proposition 1.5

Soit $X : \Omega \longrightarrow E$ une variable aléatoire sur Ω , et $A, B \subset E$. Alors

$$(X \in A) \cup (X \in B) = (X \in A \cup B), \quad (X \in A) \cap (X \in B) = (X \in A \cap B).$$

Proposition 1.6

Soit $X : \Omega \longrightarrow E$ une variable aléatoire sur Ω , et $A \subset E$. Alors

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

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 .

Remarques.

1. Quand on s'intéresse 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,

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

Remarque.

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

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

Remarques.

1. Reprenons les notations $X(\Omega) = \{x_1, \dots, x_n\}$. La loi de X est donc entièrement déterminée par la donnée de $(P(X = x_i))_{1 \leq i \leq n}$.

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 $P(X = x)$.

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

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

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

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 $\llbracket 1, n \rrbracket$ si $X(\Omega) = \llbracket 1, n \rrbracket$, et si

$$\forall k \in \llbracket 1, n \rrbracket, P(X = k) = \frac{1}{n}.$$

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

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

Remarques.

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

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 $X \hookrightarrow \mathcal{B}(p)$, 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 Bernoulli de paramètre p . Alors $P(X = 0) = 1 - p$.

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 $(X = 1)$.

Proposition 2.5

Soient X et Y des variables de Bernoulli.

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

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

Remarques.

1. On note $X \hookrightarrow \mathcal{B}(n, p)$.
2. Si $p \in]0, 1[$, alors $X(\Omega) = \llbracket 0, n \rrbracket$.
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 $(E_k)_{1 \leq k \leq n}$ des épreuves de Bernoulli indépendantes de même paramètre $p \in [0, 1]$.
 Soit X le nombre de succès. Alors $X \hookrightarrow \mathcal{B}(n, p)$.

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

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

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 suffisamment de fois une épreuve.
2. Deux v.a.r. ayant même loi ont même espérance.

Proposition 3.2

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

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

Définition 3.3 (Variable aléatoire centrée)

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

3.2 Espérance des lois usuelles**Proposition 3.4 (Espérance d'une variable indicatrice)**

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

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 .

Proposition 3.6 (Espérance d'une loi uniforme)

Soit X une variable aléatoire réelle.

1. Si $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$, alors $E(X) = \frac{n+1}{2}$.
2. Si $X \hookrightarrow \mathcal{U}(\llbracket 0, n \rrbracket)$, alors $E(X) = \frac{n}{2}$.
3. Si $X \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)$, alors $E(X) = \frac{a+b}{2}$.

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 $X \hookrightarrow \mathcal{U}(-1, 2, 3)$, 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 $E(X) = p$.

Proposition 3.8 (Espérance d'une loi binomiale)

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

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

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

Corollaire 3.10 (Espérance de $aX + b$)

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

Corollaire 3.11

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

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 connaît les espérances.

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 $P(X = 0) = 1$.

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 $E(X) \leqslant E(Y)$.

Théorème 3.15 (Inégalité de Markov)

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

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

3.5 Formule de transfert

Théorème 3.16 (Formule de transfert)

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

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

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

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

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.

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.*

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

Remarques.

1. La variance est l'espérance, *i.e.* "la valeur attendue", du carré 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

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

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

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

4.3 Variance des lois usuelles**Proposition 4.8 (Variance d'une variable indicatrice)**

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

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.

Proposition 4.10 (Variance d'une loi uniforme)

Soit X une variable aléatoire réelle qui suit la loi uniforme sur $\llbracket a, b \rrbracket$ ($a, b \in \mathbb{Z}$, $a \leq b$). Alors

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

En particulier, si $X \hookrightarrow \llbracket 1, n \rrbracket$ ($n \in \mathbb{N}$), alors $V(X) = \frac{n^2 - 1}{12}$.

Remarque.

Il faut se souvenir que $b - a + 1$ est le nombre d'éléments de l'intervalle $\llbracket a, b \rrbracket$.

Proposition 4.11 (Variance d'une loi de Bernoulli)

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

Proposition 4.12 (Variance d'une loi binomiale)

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

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 $\frac{X - E(X)}{\sigma(X)}$ est centrée réduite, appelée variable centrée réduite associée à X .

4.5 Inégalité de Bienaymé-Tchebychev

Théorème 4.16 (Inégalité de Bienaymé-Tchebychev)

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

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

Remarques.

1. On retourne 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 $(|X - E(X)| \geq \varepsilon)$. L'inégalité peut donc se réécrire

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

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

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

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

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

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

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 $\left((X = x) \cap (Y = y)\right)_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'événements, appelé système complet d'événement associé au couple (X, Y) .

Remarque.

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

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

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

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

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

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

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

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 $P(X = x, Y = y) = 0$.

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

$$\forall i \in \llbracket 1, p \rrbracket, P(X = x_i) = \sum_{j=1}^m P(X = x_i, Y = y_j),$$

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

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.

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

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

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

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

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 conditionnelle à $(Y = y)$.

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

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

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

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

On a de plus :

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

Et enfin :

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

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

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.*

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

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

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.

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 $E(XY) = E(X)E(Y)$.

Remarque.

Attention : la réciproque est fautive : si $E(XY) = E(X)E(Y)$, en général, les variables X et Y ne sont pas indépendantes. Si $X \hookrightarrow [-1, 1]$ 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

$$\text{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right).$$

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 $\text{Cov}(X, X) = V(X)$.

Proposition 7.2

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

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

Remarque.

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

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 .

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 $\text{Cov}(X, Y) = 0$.

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 covariance est nulle sont dites 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

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

Remarque.

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

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, \dots, X_n des variables aléatoires sur Ω à valeurs respectivement dans E_1, \dots, E_n . Le n -uplet de variables aléatoires (X_1, \dots, X_n) est la variable aléatoire

$$\begin{aligned} \Omega &\longrightarrow E_1 \times \dots \times E_n \\ \omega &\longmapsto (X_1(\omega), \dots, X_n(\omega)). \end{aligned}$$

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

Proposition 8.2

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

$$\left((X_1 = x_1) \cap \dots \cap (X_n = x_n) \right)_{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)}$$

forme un système complet d'événements de Ω .

Définition 8.3 (Loi conjointe, lois marginales)

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

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

Les lois marginales sont les lois de X_1, \dots, X_n .

Remarque.

C'est bien une probabilité car, comme pour les couples de variables aléatoires, elle est définie sur les événements élémentaires de $X_1(\Omega) \times \dots \times X_n(\Omega)$, 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 $i \in \llbracket 1, n \rrbracket$, notons $n_i = \text{card}(X_i(\Omega))$, et $X_i(\Omega) = \{x_{i1}, \dots, x_{in_i}\}$. On a alors pour tout $j \in \llbracket 1, n_i \rrbracket$,

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

Remarque.

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

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

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

Remarques.

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

Proposition 8.6

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

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

8.3 Somme de variables de Bernoulli indépendantes

Voici un cadre plus rigoureux pour le théorème 2.7.

Lemme 8.7

Soient X_1, \dots, X_{n+1} ($n \in \mathbb{N}^*$) des variables aléatoires réelles, mutuellement indépendantes. Alors les variables $X_1 + \dots + X_n$ et X_{n+1} sont indépendantes.

Proposition 8.8 (Variables de Bernoulli indépendantes)

Soient X_1, \dots, X_n des variables de Bernoulli de même paramètre p , mutuellement indépendantes. Alors

$$X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p).$$

Théorème 8.9 (Variance d'une somme)

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

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

Théorème 8.10 (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

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

Théorème 8.11 (Variance d'une somme de variables deux à deux indépendantes)

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

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

Proposition 8.12 (Sommes de variables de Bernoulli indépendantes)

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

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

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.