Chapitre 26 : Variables aléatoires

Table des matières

Dans tout le chapitre, on ne considérera que des espaces probabilisés **finis**, qui modélisent des expériences aléatoires n'ayant qu'un nombre fini de résultats possibles.

1 Vocabulaire des variables aléatoires

Exemple 1.1 : On lance deux dés équilibrés à 6 faces numérotées de 1 à 6.
L'univers Ω est l'ensemble des couples d'éléments de $[\![1\,;6]\!]$, sur lequel on a équiprobabilité.
On s'intéresse à la somme des deux chiffres obtenus, que l'on note X.
En termes mathématiques, X est une fonction de Ω dans $\mathbb{R}$: à chaque résultat possible $\omega \in \Omega$, qui s'écrit comme un couple $\omega = (i\,,j) \in [\![1\,;6]\!]^2$, on associe la somme des chiffres $X(\omega) = i + j$.
L'ensemble des valeurs possibles pour X est : $X(\Omega) = [\![2\,;12]\!]$ (mais ces valeurs ne sont pas équiprobables).
Une telle fonction X est appelée **variable aléatoire**.

1.1 Définition et notations

Définition 1.2 (variable aléatoire sur un espace probabilisé fini)

Soit $(\Omega\,,P)$ un espace probabilisé fini.
Une variable aléatoire sur $(\Omega\,,P)$ est une application X dont l'ensemble de départ est Ω.
Lorsque l'ensemble d'arrivée est inclus dans $\mathbb{R}$, on dit que X est une variable aléatoire réelle.

Notations : Si X est une variable aléatoire sur un espace probabilisé fini $(\Omega\,,P)$.

- L'ensemble $X(\Omega)$ (ensemble des valeurs possibles pour X) est fini (et non vide). On dit que $X(\Omega)$ est le support de X. En général, on ne précisera pas l'ensemble d'arrivée d'une variable aléatoire X, car seul l'ensemble $X(\Omega)$ des valeurs prises par X a en fait un intérêt (on pourra considérer que l'ensemble d'arrivée de X est $X(\Omega)$).
- Lorsque A est une partie de E, on note $\{X \in A\}$ ou $(X \in A)$ l'événement défini par $X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\}$ (qui est une partie de l'univers Ω), et la probabilité de cet événement est notée $P(X \in A)$.
Lorsque x est un élément de E, on adoptera des notations similaires pour des événements du type $(X = x)$, $(X \leqslant x)$, etc. ainsi que pour leurs probabilités.

Exemple 1.3 : Reprenons l'exemple du lancer de deux dés.
On a équiprobabilité sur l'univers $\Omega = [\![1\,;6]\!]^2$, de cardinal 36.

1. L'événement $(X = 4) = \{(1\,,3);(2\,,2);(3\,,1)\}$ est composé de 3 éléments, et donc : $P(X=4) = \dfrac{3}{36} = \dfrac{1}{12}$.
2. L'événement $(X \geqslant 10) = \{(4\,,6);(5\,,5);(6\,,4);(5\,,6);(6\,,5);(6\,,6)\}$ est composé de 6 éléments, et donc : $P(X \geqslant 10) = \dfrac{6}{36} = \dfrac{1}{6}$. On a aussi $P(X \geqslant 10) = P(X=10) + P(X=11) + P(X=12) = \dfrac{3}{36} + \dfrac{2}{36} + \dfrac{1}{36}$, car $\{X \geqslant 10\} = \{X=10\} \cup \{X=11\} \cup \{X=12\}$ et cette union est disjointe.

1.2 Loi d'une variable aléatoire

Exemple 1.4 : On reprend l'exemple du lancer de deux dés, pour lequel on avait pris comme univers $\Omega = [\![1\,;6]\!]^2$ avec P la loi de probabilité uniforme sur Ω, et X la variable aléatoire correspondant à la somme des deux dés, pour laquelle l'ensemble des valeurs possibles est $X(\Omega) = [\![2\,;12]\!]$.
On aurait aussi pu modéliser la situation différemment, en prenant comme univers $\Omega' = [\![2\,;12]\!]$: on aurait alors une autre loi de probabilité P' sur Ω', qui n'est plus la loi uniforme.
Cette nouvelle loi de probabilité sur l'univers $\Omega' = X(\Omega)$ est ce que nous appellerons la **loi de la variable aléatoire $\boldsymbol{X}$**, notée P_X.

Définition 1.5 (loi d'une variable aléatoire)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire sur $(\Omega\,,P)$.
On appelle loi de la variable aléatoire X la loi de probabilité P_X sur $X(\Omega)$ définie par :

$$P_X : \begin{array}{rcl} \mathcal{P}\big(X(\Omega)\big) & \longrightarrow & [0\,;1] \\ A & \longmapsto & P(X \in A) \end{array}$$

Démonstration. Montrons que P_X est bien une loi de probabilité sur l'ensemble fini $X(\Omega)$.

1. Pour tout $A \subset X(\Omega)$, $P_X(A) = P(X \in A) \in [0\,;1]$, donc P_X est bien à valeurs dans $[0\,;1]$.
2. Pour $A = X(\Omega)$, on a : $P_X(X(\Omega)) = P(X \in X(\Omega)) = P(\Omega) = 1$.
3. Soient A et B deux parties disjointes de $X(\Omega)$.
   On a alors $(X \in A \cup B) = (X \in A) \cup (X \in B)$ et ces deux ensembles sont disjoints, donc :
   $P_X(A \cup B) = P((X \in A) \cup (X \in B)) = P(X \in A) + P(X \in B) = P_X(A) + P_X(B)$, par additivité de la loi de probabilité P.

L'application P_X vérifie donc les axiomes d'une loi de probabilité. □

Remarque : On a vu qu'une loi de probabilité est déterminée de manière unique par les probabilités des événements élémentaires, c'est-à-dire des singletons.
Par conséquent, la loi P_X de la variable aléatoire X est déterminée de manière unique par la distribution de probabilité $(P_X(\{x\}))_{x \in X(\Omega)}$ que l'on note $(P(X = x))_{x \in X(\Omega)}$.

Conséquence : Pour déterminer la loi d'une variable aléatoire X sur un espace probabilisé fini $(\Omega\,, P)$:

1. on détermine son support i.e. l'ensemble $X(\Omega)$ (ensemble des valeurs prises par X) ;
2. pour chaque valeur $x \in X(\Omega)$, on calcule $P(X = x)$.

Remarque : La somme de toutes les probabilités $P(X = x)$ doit être égale à 1.

Exemple 1.6 : On reprend l'exemple de la variable aléatoire X définie dans le cadre du lancer de deux dés.
On a $X(\Omega) = [\![2\,;12]\!]$, et la loi de X est donnée par le tableau suivant :

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Notation : Pour signifier que deux variables aléatoires X_1 et X_2 ont la même loi i.e. $P_{X_1} = P_{X_2}$, on note $X_1 \sim X_2$.

Remarque : Deux variables aléatoires X_1 et X_2 ayant la même loi ne sont pas nécessairement égales.
Dans l'exemple précédent, si on appelle X_1 le résultat du premier dé et X_2 celui du second, X_1 et X_2 suivent la loi uniforme sur $[\![1\,;6]\!]$, mais ces deux variables aléatoires ne sont pas égales.
En effet, pour $\omega = (5{,}3) \in \Omega = [\![1{,}6]\!]^2$, on a $X_1(\omega) = 5 \neq X_2(\omega) = 3$ (X_1 et X_2 sont donc distinctes en tant que fonctions de Ω dans $\mathbb{R}$).

1.3 Image d'une variable aléatoire par une fonction

Définition 1.7 (image d'une variable aléatoire par une fonction)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,, P)$, et soit f une fonction définie sur $X(\Omega)$.
On peut alors définir l'application composée $f \circ X$: on obtient ainsi une nouvelle variable aléatoire sur $(\Omega\,, P)$, appelée image de X par f et notée $f(X)$.

Proposition 1.8 (loi de l'image d'une variable aléatoire)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,, P)$, et soit f une fonction définie sur $X(\Omega)$.
La loi de la variable aléatoire $f(X)$ est donnée par :

1. $f(X)(\Omega) = f\big(X(\Omega)\big)$;
2. $\forall y \in f\big(X(\Omega)\big), \quad P\big(f(X) = y\big) = \displaystyle\sum_{x \in f^{-1}(\{y\})} P(X = x)$.

Exemple 1.9 : On considère le jeu d'argent suivant : un joueur lance deux dés ; s'il fait 10 ou plus, il gagne deux euros ; sinon il perd un euro.
On garde les notations introduites précédemment pour la variable aléatoire X, égale à la somme des deux dés.
On note Y la variable aléatoire modélisant le gain du joueur. Déterminer la loi de Y.

Remarque : La loi de $f(X)$ ne dépend que de la loi de X et de la fonction f : si deux variables aléatoires X_1 et X_2 ont la même loi, alors $f(X_1)$ et $f(X_2)$ ont également la même loi.

2 Espérance et variance

2.1 Espérance d'une variable aléatoire réelle ou complexe

Dans la suite, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Définition 2.1 (espérance d'une variable aléatoire)

Soit $(\Omega\,, P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle ou complexe sur $(\Omega\,, P)$.
L'espérance de la variable aléatoire X, notée $\mathrm{E}(X)$, est le réel défini par :

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

Remarques :

1. L'espérance de X est donc la moyenne des valeurs prises par X, pondérées par leurs probabilités respectives, c'est donc un indicateur de position.
2. L'espérance de X ne dépend que de la loi P_X : si X_1 et X_2 ont même loi, alors $\mathrm{E}(X_1) = \mathrm{E}(X_2)$.
3. Si X est une variable aléatoire constante, de valeur $\lambda \in \mathbb{K}$, alors $E(X) = \lambda$.

Exemple 2.2 : Reprenons l'exemple précédent du jeu d'argent avec le lancer de deux dés.
On avait appelé Y la variable aléatoire représentant le gain. Déterminer $\mathrm{E}(Y)$.

Proposition 2.3 (espérance d'une indicatrice)

Soit $(\Omega\,, P)$ un espace probabilisé fini, et soit A un évènement de Ω.
On a la relation suivante : $E(\mathbb{1}_A) = P(A)$.

Propriétés de l'espérance :

Lemme 2.4 (expression de l'espérance à partir des probabilités élémentaires de Ω)

Soit $(\Omega\,, P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle ou complexe sur $(\Omega\,, P)$.
Son espérance est donnée par la formule :

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

Théorème 2.5 (linéarité de l'espérance)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soient X et Y deux variables aléatoires réelles ou complexes sur (Ω,P), et soit $\lambda \in \mathbb{C}$. On a la relation : $\mathrm{E}(X + \lambda Y) = \mathrm{E}(X) + \lambda\,\mathrm{E}(Y)$.

Remarque : L'espérance est donc une forme linéaire sur le $\mathbb{K}$-espace vectoriel des variables aléatoires réelles ou complexes sur (Ω,P) (c'est-à-dire $\mathcal{F}(\Omega,\mathbb{K})$).

Cas particulier : Pour X variable aléatoire réelle ou complexe et $(a\,,b) \in \mathbb{K}^2$, on a : $\mathrm{E}(aX+b) = a\,\mathrm{E}(X) + b$.

Conséquence : Soit X une variable aléatoire réelle ou complexe, dont on note m l'espérance.
Par linéarité de l'espérance : $E\big(X-m\big) = \mathrm{E}(X) - m = 0$.
Ainsi, la variable aléatoire $X - \mathrm{E}(X)$ a pour espérance 0 : on dit qu'une telle variable aléatoire est **centrée**.

Définition 2.6 (variable aléatoire centrée)

Une variable aléatoire réelle ou complexe est dite centrée lorsque son espérance est nulle.

Proposition 2.7 (inégalité triangulaire pour l'espérance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle ou complexe sur $(\Omega\,,P)$.
On a l'inéquation suivante : $|E(X)| \leqslant E(|X|)$.

Théorème 2.8 (positivité de l'espérance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle sur $(\Omega\,,P)$.
Si X est positive ($i.e.$ $X(\Omega) \subset \mathbb{R}^+$), alors $\mathrm{E}(X) \geqslant 0$.

Corollaire 2.9 (croissance de l'espérance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles sur $(\Omega\,,P)$.
Si $X \leqslant Y$, alors $\mathrm{E}(X) \leqslant \mathrm{E}(Y)$.

Théorème 2.10 (théorème de transfert)

Soit $(\Omega\,,P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,,P)$, et soit $f : X(\Omega) \to \mathbb{K}$.
Alors
$$\mathrm{E}\left(f(X)\right) = \sum_{x \in X(\Omega)} P(X=x)\,f(x)$$

Exemple 2.11 : En reprenant la variable aléatoire Y définie dans l'exemple du jeu d'argent, retrouver $\mathrm{E}(Y)$ en utilisant le théorème de transfert.

Moments d'une variable aléatoire :
Soit X une variable aléatoire réelle ou complexe sur un espace probabilisé fini $(\Omega\,,P)$ et soit $k \in \mathbb{N}$.
Le moment d'ordre k de la variable aléatoire X est l'espérance de X^k, qui est donc donnée par la formule :
$$\mathrm{E}(X^k) = \sum_{x \in X(\Omega)} P(X=x)\,x^k$$

Remarques :

1. Le moment d'ordre 0 vaut 1.
2. Le moment d'ordre 1 n'est rien d'autre que l'espérance.

2.2 Variance et écart-type d'une variable aléatoire réelle

Définition 2.12 (variance et écart-type d'une variable aléatoire réelle)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle sur $(\Omega\,,P)$.

1. La variance de la variable aléatoire X, notée $\mathrm{V}(X)$, est le réel positif défini par :

$$\mathrm{V}(X) = E\left[\left(X - E(X)\right)^2\right]$$

2. L'écart-type de la variable aléatoire X, noté $\sigma(X)$, est le réel positif défini par : $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarques :

1. La variance et l'écart-type mesurent la **dispersion** de la variable aléatoire par rapport à son espérance.
2. Une variable aléatoire constante a une variance nulle.
3. La variance et l'écart-type de X ne dépendent que de la loi P_X de la variable aléatoire X : si X_1 et X_2 ont même loi, alors $\mathrm{V}(X_1) = \mathrm{V}(X_2)$ et $\sigma(X_1) = \sigma(X_2)$.
4. Le théorème de transfert donne une première formule pour calculer la variance :

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

Théorème 2.13 (théorème de König-Huygens)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle sur $(\Omega\,,P)$.
On a l'égalité suivante :

$$\boxed{\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2}$$

Exemple 2.14 : Reprenons l'exemple du jeu d'argent avec deux dés, déterminer $\sigma(Y)$.

Propriétés de la variance et de l'écart-type :

Proposition 2.15 (variance et écart-type de $aX + b$)

Soit $(\Omega\,,P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,,P)$, et soient a et b deux réels. On a les relations :

1. $\mathrm{V}(aX + b) = a^2\,\mathrm{V}(X)$
2. $\sigma(aX + b) = |a|\,\sigma(X)$

Remarque : En général : $\mathrm{V}(X + Y) \neq \mathrm{V}(X) + \mathrm{V}(Y)$, $\sigma(X + Y) \neq \sigma(X) + \sigma(Y)$.

Conséquence : Soit X une variable aléatoire réelle, dont on note m l'espérance et σ l'écart-type.
On suppose que $\sigma > 0$.
On a déjà vu que $\mathrm{E}(X - m) = 0$ (variable aléatoire centrée), et la propriété précédente donne $\sigma(X - m) = \sigma(X) = \sigma$.
Si on divise maintenant cette variable aléatoire $X - m$ par σ, alors elle reste centrée, et son écart-type sera lui-aussi divisé par σ.
Ainsi, la variable aléatoire $\dfrac{X - m}{\sigma} = \dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est centrée et a pour écart-type 1 : on dit qu'une telle variable aléatoire est **centrée réduite**.

Définition 2.16 (variable aléatoire réduite)

Une variable aléatoire réelle est dite réduite lorsque sa variance est égale à 1 (ou, de manière équivalente, lorsque que son écart-type est égal à 1).

3 Lois usuelles

3.1 Loi uniforme

Définition 3.1 (variable aléatoire suivant une loi uniforme)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,, P)$, et soit E un ensemble fini et non vide.
On dit que la variable aléatoire X suit la loi uniforme sur E, ce que l'on note $X \sim \mathscr{U}(E)$ si :

1. $X(\Omega) = E$;
2. la loi P_X de la variable aléatoire X est la loi de probabilité uniforme sur E, *i.e.* :

$$\forall x \in E,\, P(X = x) = \frac{1}{|E|}$$

Interprétation : Cette loi traduit une situation d'**équiprobabilité** : on « choisit au hasard » une valeur de l'ensemble E, chaque valeur ayant la même probabilité d'être obtenue.

Proposition 3.2 (espérance et variance pour la loi uniforme sur $[\![1\,;n]\!]$)

Soit $n \in \mathbb{N}^*$, et soit X une variable aléatoire qui suit la loi uniforme sur $[\![1\,;n]\!]$. Alors :

$$\mathrm{E}(X) = \frac{n+1}{2} \qquad \mathrm{V}(X) = \frac{n^2-1}{12} = \frac{(n-1)(n+1)}{12}$$

3.2 Loi de Bernoulli

Définition 3.3 (variable aléatoire suivant une loi de Bernoulli)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,, P)$, et soit $p \in [0\,;1]$.
On dit que la variable aléatoire X suit la loi de Bernoulli de paramètre p, ce que l'on note $X \sim \mathscr{B}(p)$, si :

1. $X(\Omega) \subset \{0{,}1\}$;
2. $P(X = 1) = p$ et $P(X = 0) = 1 - p$.

On dit aussi que X est une variable aléatoire de Bernoulli de paramètre p.

Interprétation : Cette loi apparaît lorsqu'on s'intéresse à la réalisation ou non d'un certain événement S, appelé **succès** (on parle d'épreuve de Bernoulli pour une telle expérience aléatoire) : la variable aléatoire X qui vaut 1 si S est réalisé, 0 sinon, suit alors la loi de Bernoulli de paramètre $p = P(S)$.

Remarque : Si X suit la loi $\mathscr{B}(p)$, il en est de même de X^2, X^3, ...

Proposition 3.4 (espérance et variance pour une loi de Bernoulli)

Soit $p \in [0\,;1]$, et soit X une variable aléatoire qui suit la loi de Bernoulli $\mathscr{B}(p)$. Alors :

$$\mathrm{E}(X) = p \qquad\qquad \mathrm{V}(X) = p(1-p)$$

3.3 Loi binomiale

Définition 3.5 (variable aléatoire suivant une loi binomiale)

Soit $(\Omega\,, P)$ un espace probabilisé fini, soit X une variable aléatoire sur $(\Omega\,, P)$, et soient $n \in \mathbb{N}^*$ et $p \in [0\,;1]$.
On dit que la variable aléatoire X suit la loi binomiale de paramètres n et p, ce que l'on note $X \sim \mathscr{B}(n\,,p)$, si :

1. $X(\Omega) \subset [\![0\,;n]\!]$;
2. $\forall k \in [\![0\,;n]\!]$, $P(X=k) = \dbinom{n}{k} p^k(1-p)^{n-k}$.

Interprétation : Cette loi apparaît lorsqu'on **répète** n fois **de manière indépendante** une même épreuve de Bernoulli (ce que l'on appelle un schéma de Bernoulli), et que l'on s'intéresse au **nombre de succès**.
La variable aléatoire qui compte ce nombre de succès suit alors la loi binomiale $\mathscr{B}(n\,,p)$, où n est le nombre de répétitions et p la probabilité de succès à chaque étape.
Exemple : tirages successifs **avec remise** d'une boule dans une urne, p étant la probabilité de succès pour chaque tirage.

Remarques :

1. On a bien : $\displaystyle\sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k} = (p+1-p)^n = 1$, d'après la formule du binôme.
2. Pour $n=1$, on retrouve la loi de Bernoulli de paramètre p : $\mathscr{B}(1\,,p) = \mathscr{B}(p)$.

Théorème 3.6 (espérance et variance pour une loi binomiale)

Soient $n \in \mathbb{N}^*$ et $p \in [0\,;1]$, et soit X une variable aléatoire qui suit la loi binomiale $\mathscr{B}(n\,,p)$. Alors :

$$\mathrm{E}(X) = np \qquad\qquad \mathrm{V}(X) = np(1-p)$$

Remarque : L'espérance et la variance pour une loi binomiale $\mathscr{B}(n\,,p)$ s'obtiennent en multipliant celles d'une loi de Bernoulli $\mathscr{B}(p)$ par n, ce qui sera justifié plus loin.

4 Couples de variables aléatoires

4.1 Loi conjointe, lois marginales

Lorsque X et Y sont deux variables aléatoires sur un espace probabilisé fini (Ω,P), on peut considérer le couple $(X\,,Y)$ comme une nouvelle variable aléatoire, c'est-à-dire comme une nouvelle fonction dont l'ensemble de départ est Ω :

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

Définition 4.1 (loi conjointe et lois marginales pour un couple de variables aléatoires)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit $(X\,,Y)$ un couple de variables aléatoires sur $(\Omega\,,P)$.

1. La loi conjointe de X et Y est la loi de la variable aléatoire $(X\,,Y)$.
2. Les lois marginales du couple $(X\,,Y)$ sont les lois de X et de Y.

Remarque : La loi conjointe de X et Y est déterminée par les probabilités suivantes, pour $(x\,,y) \in X(\Omega)\times Y(\Omega)$: $P((X\,,Y) = (x\,,y)) = P(\{X = x\} \cap \{Y = y\})$, notées plus simplement $P(X = x\,,Y = y)$.

Exemple 4.2 : On lance deux dés à six faces. On appelle X la variable aléatoire égale au minimum des résultats des deux dés, Y celle égale au maximum.
On peut donner la loi conjointe de X et Y sous la forme d'un tableau explicitant les probabilités $P(X = x,Y = y)$:

x \ y	1	2	3	4	5	6
1	1/36	2/36	2/36	2/36	2/36	2/36
2	0	1/36	2/36	2/36	2/36	2/36
3	0	0	1/36	2/36	2/36	2/36
4	0	0	0	1/36	2/36	2/36
5	0	0	0	0	1/36	2/36
6	0	0	0	0	0	1/36

On peut retrouver les lois marginales d'un couple de variables aléatoires $(X\,,Y)$ à partir de la loi conjointe :

1. Les événements $\{Y = y\}$, pour $y \in Y(\Omega)$, forment un système complet d'événements, donc la formule des probabilités totales donne :
$$\forall x \in X(\Omega), \quad P(X = x) = \sum_{y \in Y(\Omega)} P(X = x\,,Y = y)$$
2. De même, les événements $\{X = x\}$, pour $x \in X(\Omega)$, forment un système complet d'événements, donc :
$$\forall y \in Y(\Omega), \quad P(Y = y) = \sum_{x \in X(\Omega)} P(X = x\,,Y = y)$$

Dans le tableau précédent, ceci revient à sommer sur la i-ème ligne pour obtenir $P(X = i)$, et sur la j-ème colonne pour obtenir $P(Y = j)$.

Remarque : En revanche, la connaissance des lois marginales ne suffit pas à déterminer la loi conjointe.
Par exemple, si on appelle X_1 (resp. X_2) la variable aléatoire égale au résultat du premier dé (resp. du second dé), alors les couples $(X_1\,,X_2)$ et $(X_1\,,X_1)$ ont les mêmes lois marginales (loi uniforme sur $[\![1\,;6]\!]$), mais pas la même loi conjointe (on a par exemple $P(X_1 = 1\,,X_2 = 2) = \frac{1}{36} \neq P(X_1 = 1\,,X_1 = 2) = 0$).

4.2 Lois conditionnelles

Définition 4.3 (loi conditionnelle d'une variable aléatoire)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit $(X\,,Y)$ un couple de variables aléatoires sur $(\Omega\,,P)$.
On considère une valeur $x \in X(\Omega)$ telle que $P(X = x) \neq 0$.
On appelle loi conditionnelle de Y sachant $\{X = x\}$ la loi de Y dans l'espace probabilisé $\big(\Omega\,,P_{\{X=x\}}\big)$.

Exemple 4.4 : On dispose de 6 urnes numérotées de 1 à 6 contenant chacune des boules blanches ou noires : l'urne numéro k contient k boules blanches et $6 - k$ boules noires.
On lance un dé à six faces, puis on tire une boule au hasard dans l'urne correspondant au numéro obtenu sur le dé. On note X la variable aléatoire correspondant à ce numéro, et Y la variable aléatoire qui vaut 1 si on obtient une boule blanche, 0 sinon. Déterminer la loi conditionnelle de Y sachant $\{X = k\}$ pour tout $k \in [\![1\,;6]\!]$.

Remarques : Supposons que l'on connaisse la loi de X et les lois conditionnelles de Y sachant $\{X = x\}$, pour tout $x \in X(\Omega)$ (en supposant que les probabilités de ces événements sont non nulles).

1. On peut alors obtenir la loi conjointe de X et Y, car $P(X = x,Y = y) = P(X = x)P(Y = y \mid X = x)$ pour tout couple $(x,y) \in X(\Omega) \times Y(\Omega)$.
2. On peut aussi obtenir la loi de Y en utilisant comme précédemment la formule des probabilités totales avec le système complet d'événements $(\{X = x\})_{x\in X(\Omega)}$:

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

Exemple 4.5 : Déterminer la loi de Y dans l'exemple précédent.

4.3 Indépendance de deux variables aléatoires

Définition 4.6 (couple de variables aléatoires indépendantes)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires sur $(\Omega\,,P)$.
On dit que X et Y sont des variables aléatoires indépendantes, ce que l'on note $X \perp\!\!\!\perp Y$, si pour toutes parties $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, c'est-à-dire :

$$P(X \in A\,,Y \in B) = P(X \in A)\,P(Y \in B)$$

Proposition 4.7 (définition équivalente de l'indépendance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires sur $(\Omega\,,P)$.
Les variables aléatoires X et Y sont indépendantes, si et seulement si

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

Reformulation : X et Y sont indépendantes si pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, la loi conditionnelle de Y sachant $\{X = x\}$ est égale à la loi de Y.
Intuitivement, ceci signifie que la connaissance de la valeur prise par X n'a aucune influence sur la valeur de Y.

Exemple 4.8 : On lance deux dés, on note X_1 la variable aléatoire égale au résultat du premier dé, X_2 celle égale au résultat du second dé. Alors les variables aléatoires X_1 et X_2 sont indépendantes.

Remarque : Si X et Y sont deux variables aléatoires **indépendantes** et que l'on connaît les lois marginales du couple $(X\,,Y)$, alors on peut déterminer la loi conjointe, car on a alors $P(X = x\,,Y = y) = P(X = x)\,P(Y = y)$ pour tout couple $(x\,,y) \in X(\Omega) \times Y(\Omega)$. Ceci est faux sans l'hypothèse d'indépendance de X et Y.

Proposition 4.9 (indépendance de deux variables aléatoires images)

Soit $(\Omega\,,P)$ un espace probabilisé fini, soient X et Y deux variables aléatoires sur $(\Omega\,,P)$, soient f une fonction définie sur $X(\Omega)$ et g une fonction définie sur $Y(\Omega)$.
Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 4.10 (espérance d'un produit de deux variables aléatoires indépendantes)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles ou complexes sur $(\Omega\,,P)$.
Si X et Y sont indépendantes, alors $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

4.4 Covariance

Définition 4.11 (covariance de deux variables aléatoires)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles sur $(\Omega\,,P)$.
La covariance des variables aléatoires X et Y, notée $\mathrm{Cov}(X\,,Y)$, est le réel défini par :

$$\mathrm{Cov}(X\,,Y) = \mathrm{E}\left[\big(X - \mathrm{E}(X)\big)\big(Y - \mathrm{E}(Y)\big)\right]$$

Propriétés de la covariance :

1. $\mathrm{Cov}(X\,,X) = V(X) \geqslant 0$
2. Symétrie : $\mathrm{Cov}(X\,,Y) = \mathrm{Cov}(Y\,,X)$
3. Bilinéarité :
   - $\mathrm{Cov}(X + \lambda X'\,,Y) = \mathrm{Cov}(X\,,Y) + \lambda\,\mathrm{Cov}(X'\,,Y)$ (linéarité à gauche)
   - $\mathrm{Cov}(X\,,Y + \lambda Y') = \mathrm{Cov}(X\,,Y) + \lambda\,\mathrm{Cov}(X\,,Y')$ (linéarité à droite)

Proposition 4.12 (formule pour la covariance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles sur $(\Omega\,,P)$.
On a l'égalité suivante :

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

Remarques :

1. Le théorème de transfert utilisé avec $f : (x,y) \mapsto xy$ permet de calculer l'espérance de XY en fonction de la loi conjointe du couple (X,Y) : $\mathrm{E}(XY) = \displaystyle\sum_{(x,y)\in(X,Y)(\Omega)} P(X = x\,,Y = y)\,xy$
2. La covariance de X et Y dépend de la loi conjointe de X et Y, elle ne peut pas être déterminée à partir de la seule connaissance des lois marginales.

Proposition 4.13 (expression de la variance d'une somme à l'aide de la covariance)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles sur $(\Omega\,,P)$.
Alors :

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

Conséquence : $\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y) \Leftrightarrow \mathrm{Cov}(X\,,Y) = 0$. Dans ce cas, on dit que X et Y sont décorrélées.

Proposition 4.14 (variance d'une somme de deux variables aléatoires indépendantes)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient X et Y deux variables aléatoires réelles sur $(\Omega\,,P)$.
Si X et Y sont indépendantes, alors X et Y décorrélées, i.e. $\mathrm{Cov}(X\,,Y) = 0$, i.e. :

$$\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$$

Remarque : Attention, la réciproque est fausse en général.

Exemple 4.15 : Soit $X \sim \mathscr{U}(\{-1;0;1\})$. Posons $Y = X^2$.

1. Déterminer la covariance de X et Y.
2. Déterminer la loi de (X,Y), puis retrouver à l'aide de celle-ci la covariance de X et Y.
3. Déterminer la loi de $\frac{X+Y}{2}$, puis retrouver à l'aide de celle-ci la covariance de X et Y.
4. Les variables X et Y sont-elles indépendantes ?

5 Famille finie de variables aléatoires

Une famille finie de variables aléatoires sur un espace probabilisé fini $(\Omega\,,P)$ est un n-uplet $(X_1\,,\dots,X_n)$ de variables aléatoires sur $(\Omega\,,P)$ (pour un certain $n\in\mathbb{N}^*$).
Les notions de loi conjointe, lois marginales et lois conditionnelles se généralisent pour de tels n-uplets.
Le théorème de transfert s'écrit, pour f une fonction de $X_1(\Omega)\times\cdots\times X_n(\Omega)$ dans $\mathbb{R}$ ou $\mathbb{C}$:

$$\mathrm{E}\left(f\big(X_1\,,\dots,X_n\big)\right)=\sum_{(x_1,\dots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)} P(X_1=x_1,\,\dots,X_n=x_n)\,f(x_1,\dots,x_n)$$

5.1 Indépendance de n variables aléatoires

Définition 5.1 (indépendance deux à deux et indépendance mutuelle)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient $X_1\,,\dots,X_n$ des variables aléatoires sur $(\Omega\,,P)$.
On dit que les variables aléatoires $X_1\,,\dots,X_n$ sont :

- deux à deux indépendantes si pour tous indices i et j distincts dans $[\![1\,;n]\!]$, X_i et X_j sont indépendantes ;
- mutuellement indépendantes si pour toutes parties $A_1\subset X_1(\Omega)\,,\dots,A_n\subset X_n(\Omega)$, les événements $\{X_1\in A_1\}\,,\dots,\{X_n\in A_n\}$ sont mutuellement indépendants c'est-à-dire :

$$\forall I\subset[\![1\,;n]\!]:\quad P\left(\bigcap_{i\in I}\{X_i\in A_i\}\right)=\prod_{i\in I}P(X_i\in A_i)$$

ou de manière équivalente,

$$\forall(x_1\,,\dots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega),\,\forall I\subset[\![1\,;n]\!]:\quad P\left(\bigcap_{i\in I}\{X_i=x_i\}\right)=\prod_{i\in I}P(X_i=x_i)$$

Exemple 5.2 : Si chaque variable aléatoire X_i décrit le résultat d'une certaine expérience aléatoire $\mathcal{E}_i$, et que ces n expériences aléatoires $\mathcal{E}_i$ sont réalisées « de manière indépendante » (par exemple : répétitions à l'identique d'une même expérience, comme un tirage de boule **avec remise**, le lancer d'un dé ou d'une pièce, etc.), alors les variables aléatoires X_i seront mutuellement indépendantes.

Remarque : Comme dans le cas des événements, l'indépendance mutuelle implique l'indépendance deux à deux, mais la réciproque est fausse.

Proposition 5.3 (lemme des coalitions)

Soit $(\Omega\,,P)$ un espace probabilisé fini, soient $X_1\,,\dots,X_n$ des variables aléatoires sur $(\Omega\,,P)$, et soient f et g deux fonctions définies sur $X_1(\Omega)\times\cdots\times X_m(\Omega)$ et sur $X_{m+1}(\Omega)\times\cdots\times X_n(\Omega)$ respectivement.
Si $X_1\,,\dots,X_n$ sont mutuellement indépendantes, alors $f(X_1\,,\dots,X_m)$ et $g(X_{m+1}\,,\dots,X_n)$ sont indépendantes.

Remarque : On peut étendre ce lemme à plusieurs coalitions, par exemple à 3 coalitions, si $X_1\,,\dots,X_n$ sont mutuellement indépendantes alors $f(X_1\,,\dots,X_m)$, $g(X_{m+1}\,,\dots,X_k)$ et $h(X_{k+1}\,,\dots,X_n)$ sont mutuellement indépendantes.

Proposition 5.4 (espérance d'un produit)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient $X_1\,,\dots,X_n$ des variables aléatoires réelles ou complexes sur $(\Omega\,,P)$.
Si $X_1\,,\dots,X_n$ sont mutuellement indépendantes, alors $\mathrm{E}(X_1\cdots X_n)=\mathrm{E}(X_1)\cdots\mathrm{E}(X_n)$.

5.2 Modélisation d'expériences aléatoires indépendantes

Théorème 5.5 (existence de variables aléatoires indépendantes suivant des lois données, admis)

Soit $n \in \mathbb{N}^*$, et soient n lois de probabilité $\mathscr{L}_1, \ldots, \mathscr{L}_n$ sur des univers finis.
Il existe un espace probabilisé fini (Ω, P) et des variables aléatoires mutuellement indépendantes $X_1, \ldots, X_n$ sur (Ω, P) telles que pour tout $i \in [\![1;n]\!]$, $X_i \sim \mathscr{L}_i$.

Conséquence : On peut modéliser n expériences aléatoires indépendantes par la donnée de n variables aléatoires mutuellement indépendantes $X_1, \cdots, X_n$ suivant des lois de probabilité données (leur existence est assurée par le théorème ci-dessus), de telle sorte que ces variables décrivent les résultats de chaque expérience.
Le fait de connaître les lois de chaque variable X_i et de savoir que ces variables sont mutuellement indépendantes sera ensuite suffisant pour mener à bien les calculs probabilistes relatifs à notre expérience, sans avoir besoin d'expliciter à quoi ressemble notre espace probabilisé fini (Ω, P).

Exemple 5.6 : On considère un schéma de Bernoulli : on répète n fois de manière indépendante une même expérience aléatoire ayant deux issues possibles, appelées succès et échec. On note $p \in [0;1]$ la probabilité supposée pour un succès.
On modélise ce schéma de Bernoulli par la donnée de n variables aléatoires mutuellement indépendantes $X_1, \cdots, X_n$ suivant la loi de Bernoulli $\mathscr{B}(p)$: la variable aléatoire X_i modélise le résultat de la i-ème expérience (X_i vaut 1 en cas de succès, 0 en cas d'échec).
Le théorème ci-dessus assure l'existence de telles variables aléatoires sur un certain espace probabilisé (Ω, P) : on dispose donc d'un cadre mathématique bien défini dans lequel on peut travailler et appliquer nos théorèmes.

Remarque : Plus généralement, le théorème précédent reste vrai pour une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires mutuellement indépendantes suivant des lois données, mais on est alors obligé de considérer un espace probabilisé infini, ce que vous apprendrez à faire l'année prochaine.
De telles considérations permettront par exemple de modéliser une « suite infinie de tirs à pile ou face ».

5.3 Somme de variables aléatoires indépendantes

Exemple 5.7 : Dans l'exemple précédent, la variable aléatoire $X_1 + \cdots + X_n$ représente le nombre de succès.

Théorème 5.8 (somme de variables aléatoires indépendantes suivant une loi de Bernoulli)

Soit (Ω, P) un espace probabilisé fini, et soient $X_1, \cdots, X_n$ des variables aléatoires sur (Ω, P).
On suppose que ces n variables aléatoires sont mutuellement indépendantes, et que chacune suit la loi de Bernoulli $\mathscr{B}(p)$, avec $p \in [0;1]$.
Alors la variable aléatoire $X_1 + \cdots + X_n$ suit la loi binomiale $\mathscr{B}(n, p)$.

Conséquence : Si on répète n fois de manière indépendante une même expérience aléatoire ayant deux issues possibles appelées succès et échec (schéma de Bernoulli), alors la variable aléatoire X qui compte le nombre de succès suit la loi binomiale $\mathscr{B}(n, p)$, où p est la probabilité de succès à chaque étape.

Démonstration du théorème 5.8.
On note X la variable aléatoire $X_1 + \cdots + X_n$. Comme chaque X_i prend comme valeurs 0 ou 1, alors $X(\Omega) \subset [\![0;n]\!]$.
Soit $k \in [\![0;n]\!]$. On veut montrer que $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
L'événement $\{X = k\}$ est réalisé lorsque k des variables X_i prennent la valeur 1 et les autres la valeur 0.
Si on note A_k l'ensemble des parties à k éléments de $[\![1;n]\!]$, on a donc :

$$\{X = k\} = \bigcup_{I \in A_k} B_I \qquad \text{où} \quad B_I = \left(\bigcap_{i \in I} \{X_i = 1\}\right) \cap \left(\bigcap_{i \in [\![1;n]\!] \setminus I} \{X_i = 0\}\right)$$

Comme les variables aléatoires X_i sont mutuellement indépendantes, on a

$$P(B_I) = \left(\prod_{i\in I} P(X_i = 1)\right) \times \left(\prod_{i\in [\![1;n]\!]\setminus I} P(X_i = 0)\right) = p^k(1-p)^{n-k}$$

Les événements B_I étant incompatibles deux à deux, on a donc, par additivité de P :

$$P(X = k) = \sum_{I\in A_k} P(B_I) = \sum_{I\in A_k} p^k(1-p)^{n-k} = \binom{n}{k} p^k(1-p)^{n-k}$$

car A_k est de cardinal $\binom{n}{k}$. La variable aléatoire X suit donc la loi binomiale $\mathscr{B}(n\,,p)$. □

Conséquence : espérance et variance pour une loi binomiale
Avec les mêmes notations, on a par linéarité de l'espérance : $\mathrm{E}(X) = \mathrm{E}(X_1+\cdots+X_n) = \mathrm{E}(X_1)+\cdots+\mathrm{E}(X_n) = np$.
On retrouve ainsi le fait que toute variable aléatoire suivant la loi binomiale $\mathscr{B}(n\,,p)$ admet pour espérance np.
On peut appliquer le même raisonnement pour retrouver la variance d'une loi binomiale, en utilisant en plus le théorème suivant.

Théorème 5.9 (variance d'une somme de variables aléatoires deux à deux indépendantes)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient $X_1\,,\cdots,X_n$ des variables aléatoires réelles sur $(\Omega\,,P)$.
Alors

$$\mathrm{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{V}(X_i) \,+\, 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i\,,X_j)$$

En particulier, si $X_1\,,\cdots,X_n$ sont décorrélées, alors $\mathrm{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{V}(X_i)$.

Démonstration. Par bilinéarité et symétrie de la covariance, on a :

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

En particulier, pour $X_1\,,\cdots,X_n$ décorrélées *i.e.* $\mathrm{Cov}(X_i\,,X_j) = 0$ où $i \neq j$, on a $\mathrm{V}(X_1+\cdots+X_n) = \sum_{i=1}^n \mathrm{V}(X_i)$. □

Remarque : Pour le cas particulier, il suffit que les X_i soient deux à deux indépendantes (ce qui est vérifié si les X_i sont mutuellement indépendantes).

6 Inégalités de concentration

Théorème 6.1 (inégalité de Markov)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle **positive** sur $(\Omega\,,P)$.

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

Remarque : L'hypothèse de positivité est primordiale. Contre-exemple : considérons la variable aléatoire X telle que $X(\Omega) = \{-2;2\}$ et $P(X=2) = P(X=-2) = \frac{1}{2}$, alors $P(X\geqslant 1) = \frac{1}{2} > \mathrm{E}(X) = 0$.

Exemple 6.2 : Les salaires étant positifs, la part de la population percevant un salaire supérieur à 5 fois le salaire moyen est au maximum d'un cinquième.

Corollaire 6.3 (inégalité de Markov généralisée)

Soit $(\Omega\,,P)$ un espace probabilisé fini, soit X une variable aléatoire réelle sur $(\Omega\,,P)$, et soit φ une fonction croissante définie sur un intervalle I à valeurs dans $\mathbb{R}_+^*$.

$$\forall a \in I, \quad P(X \geqslant a) \leqslant \frac{\mathrm{E}(\varphi(X))}{\varphi(a)}$$

Remarque : Ce corollaire n'est pas écrit au programme. Cependant, il permet dans la pratique d'obtenir des inégalités de concentration.

Théorème 6.4 (inégalité de Bienaymé-Tchebychev)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soit X une variable aléatoire réelle sur $(\Omega\,,P)$.

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

Corollaire 6.5 (i.d.c. pour une moyenne de variables aléatoires indépendantes de même loi)

Soit $(\Omega\,,P)$ un espace probabilisé fini, et soient $X, X_1\,, \cdots, X_n$ des variables aléatoires sur $(\Omega\,,P)$ mutuellement indépendantes et suivant la même loi.
Posons $M_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$. Alors

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

Remarque : En particulier, $\forall \varepsilon > 0$, $P\big(|M_n - \mathrm{E}(X)| \geqslant \varepsilon\big) \underset{n \to +\infty}{\longrightarrow} 0$. La moyenne empirique M_n fournit donc un estimateur de $\mathrm{E}(X)$.