Chapitre 37

Probabilités 2 – Variables aléatoires

Plan du chapitre

Hypothèse

Dans tout ce chapitre, $(\Omega, \mathbb{P})$ désigne un espace probabilisé fini (i.e. Ω désigne un univers fini et $\mathbb{P}$ est une probabilité définie sur Ω).
E et F sont des ensembles non vides. $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$

1 Introduction aux variables aléatoires

Le formalisme vu au chapitre précédent, à savoir un univers Ω et une probabilité $\mathbb{P}$, ne permet pas toujours de traiter efficacement un problème en probabilité. Il est fréquent qu'on s'intéresse non pas à la probabilité qu'un résultat ω se produise, mais à la probabilité qu'une fonction $f(\omega)$ prenne une certaine valeur. Par exemple, on lance trois dés à six faces et on souhaite évaluer la probabilité de l'évènement suivant :

$$A = \text{"la somme des valeurs des deux premiers dés est égale à la valeur du troisième''}$$

Dans ce cas, l'univers est $\Omega = [\![1,6]\!]^3$ que l'on munit de la probabilité uniforme $\mathbb{P}$. Un résultat ω est de la forme $(i,j,k) \in \Omega$, avec i,j,k les valeurs respectives du premier, deuxième et troisième dés. On cherche alors la probabilité que $i+j=k$, donc l'évènement A se réécrit :

$$A = \Big\{(1,1,2),\ (1,2,3),\ (2,1,3), \cdots, (3,3,6),\ (4,2,6),\ (5,1,6)\Big\}$$

Mais si on définit $f(\omega) = f(i,j,k) = i+j-k$, alors l'évènement A s'écrit plus succintement

$$A = \{\omega \in \Omega \mid f(\omega) = 0\}$$

L'introduction de cette fonction f permet donc de reformuler et d'exprimer de manière concise les événements dont on veut calculer la probabilité. On appelle cette fonction f une *variable aléatoire* et on la note en général X.

1.1 Définition d'une variable aléatoire

Définition 37.1

On dit que X est une variable aléatoire (abrégé v.a.) sur Ω si X est une application de Ω dans un certain ensemble E :

$$\begin{aligned} X : \Omega &\to E \\ \omega &\mapsto X(\omega) \end{aligned}$$

Une variable aléatoire X est donc une fonction dont la valeur dépend du résultat d'une expérience aléatoire : à chaque résultat $\omega \in \Omega$, on associe une valeur $X(\omega) \in E$.

Remarque. L'ensemble

$$X(\Omega) = \big\{X(\omega) \bigm| \omega \in \Omega\big\}$$

correspond à l'ensemble des valeurs prises par X. On dit que X est une v.a. réelle (abrégé v.a.r.) si $X(\Omega) \subset \mathbb{R}$, et complexe si $X(\Omega) \subset \mathbb{C}$.

Exemple 1. On lance deux dés à six faces, de sorte qu'on prend l'univers $\Omega = [\![1,6]\!]^2$. Pour tout $\omega = (i,j) \in \Omega$ on peut poser la variable aléatoire

$$X(\omega) = \max(i,j)$$

Autrement dit, X correspond au maximum des deux valeurs. On a alors $X(\Omega) = [\![1,6]\!]$. C'est donc une v.a.r.

Remarque. L'ensemble E de la Définition 37.1 n'a que peu d'importance. On prend en général le plus petit ensemble usuel qui contient $X(\Omega)$, par exemple $E = \mathbb{R}$ pour une v.a.r.

Malgré ce nom, une variable aléatoire X n'est ni une variable, ni aléatoire. Ce n'est pas une variable, mais une application. Elle n'est pas non plus aléatoire en soi : l'aléatoire se situe au niveau de la probabilité $\mathbb{P}$ associée à une expérience, qui à une issue ω associe sa probabilité $\mathbb{P}(\{\omega\})$.

1.2 Notations incontournables

- Soit $X : \Omega \to E$ une v.a. et A une **partie de E**. Alors on note, de deux façons possibles :

  Comme $A \subset E$ (et donc $A \not\subset \Omega$ a priori), l'ensemble A **n'est pas un événement**. Par contre, $\{X \in A\}$ **est un événement** car il est inclus dans Ω. Plus précisément $\{X \in A\} = X^{-1}(A)$.

- Comme $\{X \in A\}$ est un événement, on peut en particulier évaluer sa valeur avec une probabilité $\mathbb{P}$. Pour alléger les notations, on écrit :

  La valeur de $\mathbb{P}(X \in A)$ équivaut à la "probabilité" que $X(\omega)$ soit dans la partie A (où ω est un résultat d'une expérience aléatoire).

- On peut définir d'autres événements à partir de X : pour tout $x \in E$

$$\{X = x\} := \{\omega \in \Omega \mid X(\omega) = x\}$$

  Et si $E = \mathbb{R}$,

$$\{X \leq x\} := \{\omega \in \Omega \mid X(\omega) \leq x\}$$

$$\{X \geq x\} := \{\omega \in \Omega \mid X(\omega) \geq x\}$$

$$\{X < x\} := \{\omega \in \Omega \mid X(\omega) < x\}$$

$$\{X > x\} := \{\omega \in \Omega \mid X(\omega) > x\}$$

- On peut, là encore, évaluer la probabilité de ces événements :

$$\mathbb{P}(X = x) \qquad \mathbb{P}(X \leq x) \qquad \mathbb{P}(X > x) \qquad \text{etc.}$$

- Plus généralement si $\mathfrak{P}$ est une propriété qui est vérifiée ou non par un élémént de E, l'événement $\{X \text{ vérifie } \mathfrak{P}\}$ est défini comme l'ensemble des $\omega \in \Omega$ tels que $X(\omega)$ vérifie $\mathfrak{P}$. On peut donc écrire les événements

$$\{X \in 2\mathbb{N}\} \qquad \{X \text{ est premier}\} \qquad \left\{X^2 - X \geq 0\right\}$$

  et évaluer leur probabilité.

Exemple 2. On lance une pièce (équilibrée) 3 fois et on note X la v.a. qui correspond au nombre de "pile" obtenus. Alors (on justifiera proprement ces calculs ultérieurement) :

$$\mathbb{P}(X = 0) = \ldots \qquad\qquad \mathbb{P}(X \leq 1) = \ldots$$
$$\mathbb{P}(X = 1) = \ldots \qquad\qquad \mathbb{P}(X > 3) = \ldots$$

2 Loi d'une variable aléatoire

2.1 Définition d'une loi

Théorème 37.2

Soit $X : \Omega \to E$ une v.a. (et $\mathbb{P}$ une probabilité sur Ω). On définit l'application

$$\begin{aligned} \mathbb{P}_X = \mathbb{P}(X \in \bullet) : \mathscr{P}(E) &\to [0,1] \\ A &\mapsto \mathbb{P}(X \in A) \end{aligned}$$

L'application $\mathbb{P}_X$ est une probabilité **sur E** (et non Ω) appelée la <u>loi de X</u>.

⚠ Ne pas confondre $\mathbb{P}_X$ définie ci-dessus avec une loi conditionnelle $\mathbb{P}_B$ (où B est un évènement tel que $\mathbb{P}(B) > 0$). La notation $\mathbb{P}_X$ est officielle, mais en l'occurrence, il vaut mieux retenir la notation "$\mathbb{P}(X \in \dots)$".

Démonstration.

□

Exemple 3. Soit X une v.a. telle que $X(\Omega) = [\![1,6]\!]$. La probabilité que X soit pair est $\mathbb{P}(X \in \{2,4,6\})$. En écrivant $\{2,4,6\}$ comme la réunion disjointe des singletons $\{2\}$, $\{4\}$ et $\{6\}$, on a donc

$$\mathbb{P}(X \text{ est pair}) = \mathbb{P}(X = 2) + \mathbb{P}(X = 4) + \mathbb{P}(X = 6)$$

Comme on le voit sur cet exemple, la probabilité $\mathbb{P}(X \in A)$ se déduit des valeurs de $\mathbb{P}(X = x)$ pour tout $x \in A$. Plus généralement, comme l'application $\mathbb{P}_X$ est une probabilité sur E, elle est entièrement déterminée par les valeurs qu'elle prend sur les singletons de E, ce qui conduit au résultat suivant :

Théorème 37.3

Une loi $\mathbb{P}_X$ est entièrement déterminée par les valeurs $\mathbb{P}(X = x)$ avec $x \in E$.
De plus, on a $\sum_{x \in E} \mathbb{P}(X = x) = 1$.

Méthode

Soit X une v.a. définie sur Ω. Déterminer la loi de X consiste à :

- Identifier l'ensemble $X(\Omega)$ des valeurs prises par X.
- Déterminer pour chaque valeur $x \in X(\Omega)$ la probabilité $\mathbb{P}(X = x)$.

Exemple 4. On lance un dé à six faces et on note X la v.a. qui vaut 0 si le résultat est pair, 1 sinon.

- L'univers est $\Omega = \{1, \cdots, 6\}$: c'est l'ensemble des résultats possibles du lancer de dé.
- L'ensemble des valeurs prises par X est $X(\Omega) = \{0, 1\}$.
- Les valeurs de $\mathbb{P}(X = x)$ pour $x \in \{0, 1\}$ sont données par :

$$\mathbb{P}(X = 0) = \frac{1}{2} \qquad \text{et} \qquad \mathbb{P}(X = 1) = \frac{1}{2}$$

2.2 Lois usuelles

Remarque. Jusqu'à présent, on a toujours explicité l'univers Ω et la probabilité $\mathbb{P}$ associé à une expérience aléatoire. Dans certains cas cependant, on s'intéresse uniquement à une v.a. X, et plus précisément à sa loi, i.e. aux valeurs de $\mathbb{P}(X \in A)$ avec $A \subset E$. Dans ces cas là, on peut s'affranchir de préciser Ω et même $\mathbb{P}$! On ne précisera que la loi de X, ou, ce qui revient au même, les valeurs de $\mathbb{P}(X = x)$ pour $x \in E$.

Parfois, cela va même plus loin : on ne définit même plus X autrement que par sa loi ! Peu importe de quel univers Ω elle part. Un énoncé pourra commencer par :

Soit X une v.a. à valeurs dans $E = \ldots$ dont la loi est donnée par :

$$\forall x \in E \qquad \mathbb{P}(X = x) = \ldots$$

Cependant, pour que les valeurs de $\mathbb{P}(X = x)$ permettent de définir une loi, il faut que ces valeurs vérifient deux conditions :

Théorème 37.4

On suppose que E est fini. Soit $(z_x)_{x \in E}$ une famille de réels. On peut définir une loi pour une v.a. X en posant $\mathbb{P}(X = x) = z_x$ si et seulement si la famille (z_x) vérifie :

$$\sum_{x \in E} z_x = 1 \qquad \text{et} \qquad z_x \geq 0 \text{ pour tout } x \in E$$

On va définir un certain nombre de lois. Par le Théorème ci-dessus, il suffit pour cela de définir des valeurs pour $\mathbb{P}(X = x)$ pour tout $x \in E$ qui respectent les deux conditions du Théorème.

Définition 37.5 – Loi de Bernoulli

Soit $p \in \left[0,1\right]$. On dit qu'une v.a. X suit une loi de Bernoulli de paramètre p, si :

$$\begin{cases} X(\Omega) = \{0,1\} \\ \mathbb{P}(X=1) = p \qquad \text{et} \qquad \mathbb{P}(X=0) = 1-p \end{cases}$$

On notera $\boxed{X \sim \mathcal{B}(p)}$.

La loi de Bernoulli représente une expérience simple avec seulement deux résultats possibles :

- Le cas $X = 1$ représente un "succès", ce qui arrive avec probabilité p.
- Le cas $X = 0$ représente un "échec", ce qui arrive avec probabilité $1 - p$.

Exemple 5. On lance une pièce (équilibrée) et on note X la v.a.r. qui vaut 1 si le résultat est face, 0 si c'est pile. La v.a. X suit une loi de Bernoulli de paramètre $p = \dfrac{1}{2}$.

Exemple 6. On lance un dé (équilibré) à six faces, on note X la v.a.r. qui vaut 1 si le résultat est 5 ou plus, 0 sinon. La v.a. X suit une loi de Bernoulli de paramètre $p = \dfrac{1}{3}$.

Définition 37.6 – Loi binomiale

Soit $p \in \left[0,1\right]$ et $n \in \mathbb{N}^*$. On dit qu'une v.a. X suit une loi binomiale de paramètre (n,p), si :

$$\begin{cases} X(\Omega) = [\![0,n]\!] \\ \forall k \in [\![0,n]\!] \qquad \mathbb{P}(X=k) = \dbinom{n}{k} p^k (1-p)^{n-k} \end{cases}$$

On notera $\boxed{X \sim \mathcal{B}(n,p)}$.

Démonstration. Il faut vérifier que ceci définit bien une loi. Il faut donc s'assurer que la famille $(\mathbb{P}(X=k))_{k\in[\![0,n]\!]}$ définit bien une loi. Il est clair que $\mathbb{P}(X=k) \geq 0$. Montrons que $\sum_{k=0}^{n} \mathbb{P}(X=k) = 1$, ce qui entrainera en particulier que $\mathbb{P}(X=k) \leq 1$ et concluera. Par la formule du binôme,

$$\sum_{k=0}^{n} \mathbb{P}(X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1^n = 1$$

d'où le résultat. $\square$

Remarque. On verra ultérieurement que, si on répète n fois (de manière indépendante[1]) une expérience de Bernoulli de paramètre p, alors le nombre total de succès obtenu correspond à la loi binomiale $\mathcal{B}(n,p)$. Nous l'admettons pour le moment.

1. Ce terme sera défini ultérieurement.

Exemple 7. On lance n fois une pièce équilibrée. Si X compte le nombre total de piles obtenu, alors X suit une loi binomiale :

$$X \sim \mathcal{B}\left(n, \frac{1}{2}\right)$$

Exemple 8. On dispose d'une urne avec un nombre indéfini de boules mais une proportion $p \in \left[0,1\right]$ d'entre elles sont blanches. On tire n boules *avec remise*. Si X compte le nombre total de boules blanches tirées, alors X suit une loi binomiale :

$$X \sim \mathcal{B}(n,p)$$

Définition 37.7 – Loi uniforme

On suppose que E est fini. On dit qu'une v.a. $X : \Omega \to E$ suit une loi uniforme sur E, si

$$\begin{cases} X(\Omega) = E \\ \forall x \in E \qquad \mathbb{P}(X = x) = \dfrac{1}{\text{card}(E)} \end{cases}$$

On notera alors $\boxed{X \sim \mathcal{U}(E)}$.

Comme pour la probabilité uniforme, on en déduit immédiatement que :

$$\forall A \in \mathscr{P}(E) \qquad \mathbb{P}(X \in A) = \frac{\text{card}(A)}{\text{card}(E)}$$

Exemple 9. On lance un dé à six faces pipé, qui a deux fois plus de chances de tomber sur un résultat pair qu'un résultat impair. On lance le dé et on note X l'entier de $[\![0,2]\!]$ qui est congru modulo 3 au résultat du dé. Montrer que X suit une loi uniforme.

2.3 Compléments sur les lois

La loi d'une v.a. $X : \Omega \to E$ est une probabilité sur E. Toute loi "hérite" donc des propriétés suivantes :

Théorème 37.8

Soit $X : \Omega \to E$ une v.a. (et $\mathbb{P}$ une probabilité sur Ω). Soit A et B deux **parties de E.** Alors :

1. $\mathbb{P}(X \in E) = 1 \quad \text{et} \quad \mathbb{P}(X \in \varnothing) = 0$
2. $\mathbb{P}(X \in \overline{A}) = 1 - \mathbb{P}(X \in A)$
3. Si $A \subset B$, alors :
$$\mathbb{P}(X \in A) \leq \mathbb{P}(X \in B)$$
4. Si A et B sont disjointes, on a :
$$\mathbb{P}(X \in A \cup B) = \mathbb{P}(X \in A) + \mathbb{P}(X \in B)$$
5. Sans hypothèse sur A et B, on a :
$$\mathbb{P}(X \in A \cup B) = \mathbb{P}(X \in A) + \mathbb{P}(X \in B) - \mathbb{P}(X \in A \cap B)$$

Remarque. On a vu que si $X : \Omega \to E$ est une v.a., alors

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

Or, l'ensemble E peut être infini. Pourtant, la somme $\sum_{x \in E} \mathbb{P}(X = x)$ ci-dessus a bien un sens : en effet, l'ensemble $X(\Omega)$ est fini car $\text{card}(X(\Omega)) \leq \text{card}(\Omega)$. Ainsi, si $x \notin X(\Omega)$, on a

$$\{X = x\} = X^{-1}(\{x\}) = \varnothing \qquad \text{donc} \qquad \mathbb{P}(X = x) = 0$$

Au final, on peut donc réécrire

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

et cette dernière somme ne fait intervenir qu'un nombre fini de termes, donc a toujours un sens.

Remarque. La famille $(\mathbb{P}(X=x))_{x\in E}$ détermine entièrement $\mathbb{P}_X$, mais pas X : on peut en effet trouver des v.a. différentes X_1 et X_2 telles que $\mathbb{P}_{X_1}=\mathbb{P}_{X_2}$, cf exemple 11.

Définition 37.9

Soit A **un événement de Ω**. Alors on note $\mathbb{1}_A$ la variable aléatoire réelle suivante :

$$\mathbb{1}_A : \Omega \to \{0,1\}$$
$$x \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Cette v.a.r. est appelée indicatrice de A.

Exemple 10. Soit A **un événement** (donc $A\subset\Omega$). On considère la v.a.r. $X=\mathbb{1}_A$. Alors

$$X(\Omega) = \dots\dots$$

$$\{\mathbb{1}_A = 1\} = \dots\dots\dots\dots \qquad \{\mathbb{1}_A = 0\} = \dots\dots\dots\dots$$

Donc :

$$\mathbb{P}(X=1) = \mathbb{P}(\dots\dots) \qquad \mathbb{P}(X=0) = \mathbb{P}(\dots\dots)$$

Finalement, $X \sim \dots\dots$

3 Opérations sur les v.a.

3.1 V.a. de même loi

Définition 37.10

Soit X et Y deux v.a. à valeurs dans le même ensemble E. On dit que X et Y suivent la même loi si $\mathbb{P}_X=\mathbb{P}_Y$, càd

$$\forall A \in \mathscr{P}(E) \qquad \mathbb{P}(X\in A) = \mathbb{P}(Y\in A)$$

ou encore, ce qui est équivalent :

$$\forall x \in E \qquad \mathbb{P}(X=x) = \mathbb{P}(Y=x)$$

Notation. Si X et Y suivent la même loi, on note $\boxed{X\sim Y}$.

Théorème 37.11

La relation $\sim$ entre variables aléatoires est une relation d'équivalence.

⚠ $X \sim Y$ ne signifie pas $X = Y$!

Exemple 11. Soit X et Y deux v.a. telles que $X \sim \mathcal{B}\left(\frac{1}{2}\right)$ et $Y = 1 - X$. Montrer que X et Y ont même loi. A-t-on $X = Y$?

3.2 Opérations sur les v.a.

Si X et Y sont des v.a. réelles ou complexes définies sur Ω, alors ce sont des applications de Ω dans $\mathbb{K}$, donc des éléments de $\mathcal{F}(\Omega, \mathbb{K}) = \mathbb{K}^\Omega$. On a vu en particulier que $\mathcal{F}(\Omega, \mathbb{K})$ possède une structure d'e.v. et d'anneau, ce qui, pour tous $f, g \in \mathcal{F}(\Omega, \mathbb{K})$, donne un sens aux applications $f + g$, λf et fg. De même, on peut définir d'autres v.a.r. à partir de X, Y :

$$X + Y : \omega \mapsto X(\omega) + Y(\omega)$$

$$\text{Pour tout } \lambda \in \mathbb{K} \qquad \lambda X : \omega \mapsto \lambda X(\omega)$$

$$XY : \omega \mapsto X(\omega)Y(\omega)$$

Exemple 12. On lance deux dés à six faces et on note X et Y les résultats des deux dés. Les v.a. $2X$ et $X + Y$ ont-elles la même loi ?

Définition 37.12 – Composition

Soit $X : \Omega \to E$ une v.a. et $f : E \to F$ une application. Alors on note $f(X)$ la v.a. $f \circ X$ définie sur Ω à valeurs dans F :

$$\begin{aligned} f(X) : \Omega &\to F \\ \omega &\mapsto f(X(\omega)) \end{aligned}$$

On peut donc, sous réserve que cela ait un sens, considérer les v.a. $\sqrt{X}$, $\ln X$, e^X, etc.

La v.a. $f(X)$ est une v.a. de Ω dans F. Pour déterminer la loi de $f(X)$, il suffit de connaitre les valeurs de $\mathbb{P}(f(X)=y)$ lorsque y parcourt F. En théorie, on a le résultat suivant :

$$\forall y \in F \qquad \mathbb{P}(f(X)=y) = \sum_{x\in E,\, f(x)=y} \mathbb{P}(X=x)$$

mais cette formule se comprend mieux sur un cas pratique :

Exemple 13. Soit X une v.a. telle que $X \sim \mathcal{U}(\llbracket -2,2 \rrbracket)$. Déterminer la loi de X^2.

4 Variables aléatoires indépendantes

4.1 Définition

Notation. Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. définies sur le même univers Ω. Soit $A \subset E$ et $B \subset F$. On rappelle que $\{X \in A\}$ et $\{Y \in B\}$ sont des événements, donc des sous-ensembles de Ω. On introduit la notation :

$$\mathbb{P}(X \in A, Y \in B) := \mathbb{P}\left(\{X \in A\} \cap \{Y \in B\}\right)$$

C'est donc la probabilité que X soit dans A **ET** Y soit dans B. De plus, pour tous $x \in E$ et $y \in F$, on pose également :

$$\mathbb{P}(X = x, Y = y) := \mathbb{P}\left(\{X = x\} \cap \{Y = y\}\right)$$

C'est donc la probabilité que X est égal à x **ET** Y est égal à y.

Définition 37.13

Deux v.a. $X : \Omega \to E$ et $Y : \Omega \to F$ sont dites <u>indépendantes</u> si

Notation. Lorsque X et Y sont indépendantes, on note $\boxed{X \perp\!\!\!\perp Y}$.

Théorème 37.14 – Caractérisation de l'indépendance

Deux v.a. $X : \Omega \to E$ et $Y : \Omega \to F$ sont indépendantes si et seulement si

$$\forall x \in E \quad \forall y \in F \qquad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Démonstration.

Sens direct : on utilise la définition avec $A = \{x\}$ et $B = \{y\}$.

Sens réciproque : on suppose que

$$\forall x \in E \quad \forall y \in F \qquad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Montrons que $X \perp\!\!\!\perp Y$. Soit $A \subset E$ et $B \subset F$.

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \mathbb{P}(\{X \in A\} \cap \{Y \in B\}) \\ &= \mathbb{P}\left(\left(\bigcup_{x \in A} \{X = x\}\right) \bigcap \left(\bigcup_{y \in B} \{Y = y\}\right)\right) \\ &= \mathbb{P}\left(\bigcup_{x \in A} \bigcup_{y \in B} (\{X = x\} \cap \{Y = y\})\right)\end{aligned}$$

Or, les évènements $(\{X = x\} \cap \{Y = y\})_{x \in A, y \in B}$ sont disjoints deux à deux. Par additivité :

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

□

Corollaire 37.15

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. indépendantes. Alors pour tout $x \in E$ et pour tout $y \in F$, si $\mathbb{P}(Y = y) > 0$, on a

$$\mathbb{P}(X = x \mid Y = y) = \mathbb{P}(X = x)$$

Démonstration. Cela découle du fait que les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, cf chapitre précédent. □

Remarque. Dire que deux v.a. sont indépendantes signifient "grosso modo" que les valeurs prises par l'une n'ont aucune incidence (en termes de probabilités) sur les valeurs de l'autre.

Bien souvent, un énoncé sous-entend des situations où les v.a. sont indépendantes : différents lancers de dés ou de pièces, des tirages *avec remise* de boules dans une urne, etc.

Exemple 14. On lance deux dés à 6 faces. On note X_1 et X_2 les v.a. correspondant aux résultats de chaque lancer. Alors X_1 et X_2 sont des v.a. indépendantes.

Exemple 15. On reprend l'exemple précédent. On pose

$$U = \min(X_1, X_2) \qquad \text{et} \qquad V = \max(X_1, X_2)$$

alors U et V sont aussi des v.a. mais on va montrer qu'elles ne sont pas indépendantes.

Remarque. C'est naturel pour X_1 et X_2 de l'exemple ci-dessus, mais ce n'est pas le cas pour U et V : en effet, on doit toujours avoir $U \leq V$: les valeurs prises par U influencent donc les valeurs possibles pour V et réciproquement.

4.2 Indépendance de n variables aléatoires

Définition 37.16

Soit $X_1, \cdots, X_n$ des v.a. à valeurs dans des ensembles $E_1, \cdots, E_n$ respectivement. Les v.a. $X_1, \cdots, X_n$ sont dites (mutuellement) indépendantes si

$$\forall (x_1, \cdots, x_n) \in E_1 \times \cdots \times E_n \qquad \mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \prod_{k=1}^{n} \mathbb{P}(X_k = x_k)$$

Cette définition équivaut en fait à

$$\forall (A_1, \cdots, A_n) \in \mathscr{P}(E_1) \times \cdots \times \mathscr{P}(E_n) \qquad \mathbb{P}(X_1 \in A_1, \cdots, X_n \in A_n) = \prod_{k=1}^{n} \mathbb{P}(X_k \in A_k)$$

Enfin, on peut également définir la notion de v.a. indépendantes deux à deux :

Définition 37.17

Soit $X_1, \cdots, X_n$ des v.a. à valeurs dans des ensembles $E_1, \cdots, E_n$ respectivement. Les v.a. $X_1, \cdots, X_n$ sont dites indépendantes deux à deux si pour tous $i, j \in [\![1, n]\!]$, si $i \neq j$ alors les v.a. X_i et X_j sont indépendantes.

Comme pour les événements, l'indépendance mutuelle de v.a. est une condition plus forte que l'indépendance deux à deux d'une famille de v.a.

Théorème 37.18

Soit $X_1, \cdots, X_n$ des v.a. (mutuellement) **indépendantes**. On suppose que $X_1, \cdots, X_n$ suivent toutes la même loi de Bernoulli de paramètre $p \in [0, 1]$, ce qu'on peut noter $X_1 \sim \cdots \sim X_n \sim \mathcal{B}(p)$.

Alors la v.a. $Y = X_1 + \ldots + X_n$ suit une loi binomiale de paramètre (n, p), i.e. $Y \sim \mathcal{B}(n, p)$.

Preuve "moche" mais compréhensible. Comme $X_1, \cdots, X_n$ sont à valeurs dans $\{0, 1\}$, la v.a. Y est à valeurs dans $[\![0, n]\!]$. Soit $k \in [\![0, n]\!]$, calculons $\mathbb{P}(Y = k)$. L'événement $\{Y = k\}$ est réalisé quand exactement k v.a. parmi $X_1, \cdots, X_n$ valent 1.

Supposons que les k variables qui valent 1 soient les k premières, à savoir $X_1, \cdots, X_k$. Alors par indépendance,

$$\begin{aligned}\mathbb{P}(X_1 = 1, \cdots, X_k = 1,\ X_{k+1} = 0, \cdots, X_n = 0) &= \prod_{i=1}^{k} \mathbb{P}(X_i = 1) \prod_{i=k+1}^{n} \mathbb{P}(X_i = 0) \\ &= \prod_{i=1}^{k} p \prod_{i=k+1}^{n} (1-p) = p^k (1-p)^{n-k}\end{aligned}$$

On peut vérifier que peu importe le choix des k variables parmi $X_1, \cdots, X_n$, on retombe sur la même valeur $p^k(1-p)^{n-k}$. De plus, les événements obtenus par deux choix différents sont incompatibles. Cela permet d'écrire

$$\begin{aligned}\mathbb{P}(Y = k) &= \underbrace{\mathbb{P}(\ldots\ldots\ldots\ldots)}_{\text{un choix possible de } k \text{ v.a. mises à 1, les autres à 0}} + \underbrace{\mathbb{P}(\ldots\ldots\ldots\ldots)}_{\text{un autre choix}} + \ldots \\ &= p^k(1-p)^{n-k} + p^k(1-p)^{n-k} + \ldots\end{aligned}$$

Enfin, il y a autant de choix possibles que de façons de choisir k v.a. parmi n. On somme donc $\binom{n}{k}$ termes :

$$\mathbb{P}(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

□

Exemple 16. Si on lance n fois une pièce (équilibrée), et qu'on compte le nombre de piles obtenus, alors la v.a. correspondante suit une loi binomiale $\mathcal{B}\left(n, \frac{1}{2}\right)$. *L'indépendance des lancers de pièces est ici sous-entendue.*

4.3 L'indépendance de v.a. est conservée lorsqu'on leur applique des fonctions

Théorème 37.19

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. Soit f et g des fonctions définies sur E et F respectivement. Alors

$$X \perp\!\!\!\perp Y \implies f(X) \perp\!\!\!\perp g(Y)$$

Exemple 17. On reprend l'exemple où on lance deux dés à six faces et où l'on note X_1 et X_2 les résultats obtenus. Alors, par exemple, les v.a. X_1^2 et $\sqrt{X_2}$ sont indépendantes. Les valeurs prises par X_1^2 n'ont aucune influence sur les valeurs prises par $\sqrt{X_2}$ et réciproquement.

Théorème 37.20

Soit $X_1, \cdots, X_n$ des v.a. à valeurs dans des ensembles $E_1, \cdots, E_n$ respectivement. Soit $f_1, \cdots, f_n$ des fonctions définies sur $E_1, \cdots, E_n$ respectivement. Si $X_1, \cdots, X_n$ sont (mutuellement) indépendantes, alors les v.a.

$$f_1(X_1), \cdots, f_n(X_n)$$

sont également (mutuellement) indépendantes.

Exemple 18. Soit X, Y, Z, T quatre v.a.r. indépendantes. Alors

$$X^2, \quad e^Y, \quad \ln(1 + |Z|), \quad \arctan(T)$$

sont des v.a.r. indépendantes

Lemme 37.21 – Lemme des coalitions

Soit $X_1, \cdots, X_n$ des v.a. à valeurs dans des ensembles $E_1, \cdots, E_n$ respectivement. Soit $p_1, \cdots, p_r \in [\![1, n]\!]$ tels que $1 < p_1 < p_2 < \cdots < p_r < n$.

- Soit f_1 une fonction définie sur $E_1 \times \cdots \times E_{p_1}$, de sorte que $f_1(X_1, \cdots, X_{p_1})$ ait un sens.
- Soit f_2 une fonction définie sur $E_{p_1+1} \times \cdots \times E_{p_2}$, de sorte que $f_2(X_{p_1+1}, \cdots, X_{p_2})$ ait un sens.
- ...
- Soit f_{r+1} une fonction définie sur $E_{p_r+1} \times \cdots \times E_n$, de sorte que $f_{r+1}(X_{p_r+1}, \cdots, X_n)$ ait un sens.

Alors les v.a.

$$f_1(X_1, \cdots, X_{p_1}), \qquad f_2(X_{p_1+1}, \cdots, X_{p_2}), \qquad \cdots \qquad f_{r+1}(X_{p_r+1}, \cdots, X_n)$$

sont (mutuellement) indépendantes.

Exemple 19. Soit X, Y, Z, T quatre v.a.r. indépendantes. Alors

$$e^X Y^2 \qquad \text{et} \qquad \arctan(ZT) \qquad \text{sont indépendantes}$$

$$\sqrt{|X|} \qquad \text{et} \qquad \cos(Y^Z)\sin(T) \qquad \text{sont indépendantes}$$

5 Conditionnement et v.a.

5.1 Loi conditionnelle

Rappel : si $X : \Omega \to E$ est une v.a., alors la loi de X est l'application

$$\begin{aligned} \mathbb{P}(X \in \bullet) : \mathscr{P}(E) &\to [0, 1] \\ A &\mapsto \mathbb{P}(X \in A) \end{aligned}$$

Définition 37.22

Soit $X : \Omega \to E$ une v.a. sur l'espace probabilisé $(\Omega, \mathbb{P})$. Pour tout événement B de Ω tel que $\mathbb{P}(B) > 0$ on définit la <u>loi (conditionnelle) de X sachant B</u> comme étant l'application

$$\begin{aligned} \mathbb{P}(X \in \bullet \mid B) : \quad \mathscr{P}(E) &\to [0, 1] \\ A &\mapsto \mathbb{P}(X \in A \mid B) \end{aligned}$$

Dans la définition ci-dessus, on a $B \subset \Omega$ mais $A \subset E$.

Remarque. On a $\mathbb{P}(X \in A \mid B) = \mathbb{P}_B(X \in A)$. Autrement dit, cela correspond à la loi de X dans l'espace probabilisé $(\Omega, \mathbb{P}_B)$.

Contrairement à la loi de X, i.e. $\mathbb{P}_X$, il n'y a pas de notation spécifique pour la loi conditionnelle de X sachant B, même si $\mathbb{P}(X \in \bullet \mid B)$ est bien pratique... Comme toutes les lois, $\mathbb{P}(X \in \bullet \mid B)$ est entièrement déterminée par sa valeur sur les singletons $\{x\}$ avec $x \in E$, càd par la distribution de probabilités

$$(\mathbb{P}(X = x \mid B))_{x \in E}$$

Exemple 20. On lance un dé à six faces. On pose X la v.a. égale au numéro obtenu, et B l'événement $\{$le résultat du dé est pair$\}$. Déterminer la loi conditionnelle de X sachant B.

5.2 Conditionnement d'une v.a. par rapport à une autre

Lemme 37.23

Soit X une v.a. sur Ω. Alors la famille d'évènements $\Big((X=x)\Big)_{x\in E}$ forme un S.C.E. de Ω.

Démonstration. En effet, avec $x_1, x_2 \in E$ distincts, les évènements $(X=x_1)$ et $(X=x_2)$ sont incompatibles (aucun $\omega \in \Omega$ ne vérifie simultanément $X(\omega)=x_1$ et $X(\omega)=x_2$). De plus,

$$\bigcup_{x\in E}(X=x)=\bigcup_{x\in E}(X\in\{x\})=\left(X\in\bigcup_{x\in E}\{x\}\right)=(X\in E)=\Omega$$

d'où le résultat voulu. □

Théorème 37.24

Soit X et Y des v.a. sur Ω à valeurs dans E et F respectivement. Alors pour tout $x\in E$

$$\mathbb{P}(X=x)=\sum_{y\in F}\mathbb{P}(X=x\mid Y=y)\mathbb{P}(Y=y)$$

avec la convention $\quad \mathbb{P}(\bullet\mid Y=y)\mathbb{P}(Y=y)=0 \quad$ si $\mathbb{P}(Y=y)=0$.

Démonstration.

□

6 Couples de v.a.

6.1 Définition, loi conjointe

Définition 37.25

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. Alors on définit l'application notée (X,Y) définie par :

$$\begin{aligned}(X,Y) : \Omega &\to E \times F\\ \omega &\mapsto (X(\omega), Y(\omega))\end{aligned}$$

L'application (X,Y) est appelée un couple de v.a.

Définition 37.26 – Loi conjointe

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. La loi conjointe du couple (X,Y) est la loi du couple (X,Y). C'est donc l'application

$$\begin{aligned}\mathbb{P}_{(X,Y)} : \mathscr{P}(E \times F) &\to \big[0,1\big]\\ S &\mapsto \mathbb{P}(\ (X,Y) \in S\ )\end{aligned}$$

On a donc $S \subset E \times F$ dans la définition ci-dessus. S représente un ensemble de valeurs que peut prendre le couple (X,Y), par exemple si $E = F = \{0,1,2\}$, on peut considérer $S = \{(0,1),(1,1),(1,2)\}$. Dans ce cas, $\mathbb{P}(\ (X,Y) \in S\ )$ représente la probabilité que le couple (X,Y) prenne les valeurs $(0,1)$, $(1,1)$ ou $(1,2)$.

Théorème 37.27

La loi conjointe de (X,Y) est entièrement déterminée par la famille

$$(\mathbb{P}(X = x,\ Y = y))_{x \in E,\ y \in F}$$

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

Pour déterminer une loi conjointe d'un couple (X,Y), il suffit donc de déterminer les valeurs de $\mathbb{P}(X = x,\ Y = y)$ pour tous $x \in E$ et $y \in F$.

Exemple 21. On lance deux dés à 4 faces et on note X_1 et X_2 les v.a. correspondant aux résultats de chaque lancer. Enfin, on note

$$U = \min(X_1, X_2) \qquad V = \max(X_1, X_2)$$

Déterminer la loi conjointe de (U,V).

On peut synthétiser les résultats sous la forme d'un tableau :

$\mathbb{P}(U = \dots, V = \dots)$	$V = 1$	$V = 2$	$V = 3$	$V = 4$
$U = 1$				
$U = 2$				
$U = 3$				
$U = 4$				

Remarque. Comme on doit avoir $\sum\limits_{i \in U(\Omega)} \sum\limits_{j \in V(\Omega)} \mathbb{P}(U = i, V = j) = 1$, il faut que la somme de toutes les cases de ce tableau fasse 1.

6.2 Lois marginales

Définition 37.28

Pour tout couple (X, Y) de v.a., la loi de X et la loi de Y sont appelées des lois marginales du couple (X, Y).

Ainsi pour un couple (X, Y), on dispose de 3 lois :

1. La loi conjointe $\mathbb{P}_{(X,Y)}$ déterminée par les valeurs de $\mathbb{P}(X = x, Y = y)$.
2. La loi marginale $\mathbb{P}_X$ déterminée par les valeurs de $\mathbb{P}(X = x)$.
3. La loi marginale $\mathbb{P}_Y$ déterminée par les valeurs de $\mathbb{P}(Y = y)$.

À partir de la loi conjointe, on peut déterminer les lois marginales :

Théorème 37.29

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux v.a. Alors

$$\forall x \in E \qquad \mathbb{P}(X = x) = \sum_{y \in F} \mathbb{P}(X = x, Y = y)$$

Exemple 22. Déterminons les lois marginales de l'exemple précédent. Pour cela on reprend le tableau précédent : il suffit de faire la somme sur chaque ligne et chaque colonne pour avoir les lois marginales.

$\mathbb{P}(U=\dots,V=\dots)$	$V=1$	$V=2$	$V=3$	$V=4$	$\mathbb{P}(U=...)$
$U=1$	$\frac{1}{16}$	0	0	0	
$U=2$	$\frac{1}{8}$	$\frac{1}{16}$	0	0	
$U=3$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	0	
$U=4$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	
$\mathbb{P}(V=\dots)$					Total : 1

Remarque. Comme on l'a dit, la connaissance de la loi conjointe suffit à déterminer les lois marginales. La réciproque est fausse : par exemple ci-dessus, si on ne connait que les lois marginales, on a accès à la somme de chaque ligne et chaque colonne, i.e. 8 informations. Cela ne suffit pas à "reconstruire" les 16 valeurs du tableau qui correspondent à la loi conjointe !

6.3 Couple de deux v.a. indépendantes

Lorsque deux v.a. $X:\Omega\to E$ et $Y:\Omega\to F$ sont **indépendantes**, alors exceptionnellement les lois marginales permettent de déduire la loi conjointe. En effet, on a pour tous $x\in E$ et $y\in F$:

$$\mathbb{P}(X=x,Y=y)=\mathbb{P}(X=x)\mathbb{P}(Y=y)$$

Dans ce cas :

- si on connait la loi marginale de X, i.e. les valeurs $\mathbb{P}(X=x)$ pour tout $x\in E$, et si on connait la loi marginale de Y, i.e. les valeurs $\mathbb{P}(Y=y)$ pour tout $y\in F$ ...
- alors on connait la loi conjointe de (X,Y), i.e. les valeurs $\mathbb{P}(X=x,Y=y)$ pour tous $x\in E$ et $y\in F$.

7 Méthodes pour les exercices

Méthode

Pour déterminer la loi d'une v.a. $X : \Omega \to E$, il faut :

1. Identifier l'ensemble des valeurs prises par X, i.e. $X(\Omega)$.
2. Pour chaque $x \in X(\Omega)$, déterminer $\mathbb{P}(X = x)$.

Si en revanche la loi de X est déjà connue, il n'est en général pas nécessaire d'expliciter l'univers Ω et la probabilité $\mathbb{P}$.

Méthode

Pour déterminer la loi (conjointe) d'un couple (X, Y), on peut :

- Si X et Y sont indépendantes, déterminer les lois marginales de X et de Y en calculant $\mathbb{P}(X = x)$ et $\mathbb{P}(Y = y)$ indépendamment.
- Sinon, calculer $\mathbb{P}(X = x, Y = y)$, éventuellement en faisant apparaitre un conditionnement.