Chapitre 1

Probabilité finie

La théorie des probabilités fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec certitude.

Expérience	Univers Ω (ensemble des issues)	cardinal
Lancer d'un dé	Un entier $k \in [\![1, 6]\!]$	fini
Prélèvement de n objets en sortie d'une chaîne de production dans l'échantillon	Nombre d'objets défectueux, $k \in [\![0, n]\!]$	fini
Questionnaire à 100 questions binaires	Suite ω de 100 réponses $\omega \in \{0, 1\}^{100}$	fini
Lancer d'une pièce jusqu'à la première obtention de pile	Un entier $k \in \mathbb{N}$: le temps d'attente du premier succès	dénombrable
Temps d'attente pour une hotline	un temps $\omega \in [0, +\infty[$	continue
Mouvement d'un grain de pollen dans un liquide	Une fonction continue : la trajectoire $f : \mathbb{R}^+ \to \mathbb{R}^2$	continue
Mouvement d'un cours boursier	Une fonction continue : $f : \mathbb{R}^+ \to \mathbb{R}$	continue

Bien que le résultat précis de chacune de ces expériences soit imprévisible, l'observation et l'intuition nous amènent à penser que ces phénomènes obéissent à certaines lois. Par exemple si on jette 6000 fois le dé, on s'attend à ce que le nombre d'apparitions de la face « 3 » soit voisin de 1000. Si on met en service 100 ampoules, leurs durées de vie observées seront concentrées autour d'une certaine valeur moyenne.
La **théorie des probabilités** modélise l'expérience aléatoire dans un cadre formel en quantifiant le sentiment d'incertitude vis-à-vis d'un événement. La **statistique** permet de confronter les modèles probabilistes avec la réalité observée afin de les valider ou de les invalider. Par exemple si quelqu'un a 60 bonnes réponses sur 100 au questionnaire, est-il légitime de considérer qu'il a « mieux fait » que le hasard ?
Dans le cadre de cours, on limite l'univers à un ensemble **fini**, c'est à dire que l'on peut effectuer une énumération finie des issues de l'expérience aléatoire.

I Variable aléatoire

Exemple 1 (*Jeu*)

Pour attirer les clients, un casino propose un nouveau jeu : le croupier lance simultanément 2 dés et calcule leur somme,
- si la somme est égale à 2 ou 12, le joueur gagne 2 euros,
- si la somme est égale à 7, le casino gagne 1 euro,
- dans les autres cas, c'est nul (le joueur gagne 0 euro).

A votre avis, ce jeu est-il favorable au joueur ou au casino ? A chaque partie quelle est le gain moyen du joueur ?

Dans ce jeu, on fait intervenir le hasard en observant la somme des points marqués par deux dés. Considérons le jet d'un dé bleu et d'un dé rouge et notons S la somme des points obtenus. On modélise cette expérience en prenant l'équiprobabilité sur l'univers

$$\Omega = \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du dé bleu}} \times \overbrace{\{1,2,3,4,5,6\}}^{\text{issues du dé rouge}}.$$

Une issue, ω, est un couple (b,r) où b désigne le chiffre du dé bleu et r celui du rouge. La somme S est l'application :

$$S \left| \begin{array}{rcl} \Omega & \longrightarrow & \{2,3,\dots,12\} \\ (b,r) & \longmapsto & b+r \end{array} \right.$$

Cette application est représentée par ce tableau :

b \ r	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
2	5	6	7	8	9	10
1	6	7	8	9	10	11
2	7	8	9	10	11	12

On dit que S est une variable aléatoire sur $\{2,3,\dots,12\}$.
En fait, l'observation qui nous intéresse dans cette expérience, ce n'est pas ω, mais seulement $S(\omega)$. On aimerait connaître la probabilité que la somme des points prenne une valeur donnée, soit $P(S=k)$ pour k entier fixé entre 2 et 12.
La méthode est de déterminer l'ensemble des issues donnant $(S=k)$ en lisant le tableau ci-dessous puis de calculer la probabilité de cette ensemble. Par exemple, on a $(S=3) = \{(1,2),(2,1)\}$, puis $P(S=3) = P(\{(1,2),(2,1)\}) \overset{\text{équiprobabilité}}{\overbrace{=}} \frac{\text{card}(\{(1,2),(2,1)\})}{\text{card}(\Omega)} = \frac{2}{36}$.
On obtient ainsi :

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

Cela revient à considérer un nouvel univers :

$$\Omega' = \{2,3,4,5,6,7,8,9,10,11,12\}$$

et à munir cet ensemble de la probabilité P_S définie par le tableau des $P(S=k)$. Cette nouvelle probabilité s'appelle loi de la variable aléatoire S.

A Définition

Définition 1 (*Variable aléatoire*)

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini.
Une **variable aléatoire réelle** est une application :

$$X \left| \begin{array}{rcl} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & X(\omega) \end{array} \right. .$$

On note $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ l'ensemble des images de la variable aléatoire.
Comme Ω est fini, $X(\Omega)$ est aussi finie. On s'autorise désormais à écrire l'ensemble des images

sous la forme $X(\Omega) = \{x_1, x_2, \dots, x_n\}$.

Remarque 1

- Malgré son nom, une variable aléatoire n'est pas une variable (c'est une fonction) et elle n'est pas aléatoire.
- Il est toujours possible de quantifier les valeurs prises par une variable aléatoire. Par exemple si X représente le sexe de l'enfant à la naissance, on posera le label 1 pour le sexe féminin et le label 0 pour le sexe masculin et dans ce cas $X(\Omega) = \{0, 1\}$.
- Dans ce cours l'espace de probabilité est finie et la variable aléatoire réelle, on écrira "Soit X une variable aléatoire" au lieu de "Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), P)$."

Exemple 2 (*Variable aléatoire indicatrice de A*)

Soit A un événement de Ω.
La variable aléatoire indicatrice de A, noté 1_A, est une fonction explicitant l'appartenance ou non à un événement A de toute issue de Ω :

$$1_A \left| \begin{array}{lcl} \Omega & \longrightarrow & \{0,1\} \\ \omega & \longmapsto & \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \end{array} \right. .$$

Définition 2 (*Événements associés à une variable aléatoire*)

Soit X une variable aléatoire.
Pour tout $A \subset \mathbb{R}$, on définit l'événement $\{X \in A\}$ comme étant

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}.$$

Dans le cadre de ce chapitre, nous aurons recours aux notations suivantes, pour $x \in \mathbb{R}$ et $A \subset \mathbb{R}$:

$$\begin{aligned}
(X \in A) &= X^{-1}(A) = \{\omega \in \Omega | X(\omega) \in A\} \\
(X = x) &= X^{-1}(\{x\}) = \{\omega \in \Omega | X(\omega) = x\} \\
(X \leqslant x) &= X^{-1}(]-\infty, x]) = \{\omega \in \Omega | X(\omega) \leqslant x\} \\
(X < x) &= X^{-1}(]-\infty, x[) = \{\omega \in \Omega | X(\omega) < x\} \\
(X \geqslant x) &= X^{-1}([x, +\infty[) = \{\omega \in \Omega | X(\omega) \geqslant x\} \\
(X > x) &= X^{-1}(]x, +\infty[) = \{\omega \in \Omega | X(\omega) > x\}
\end{aligned}$$

Remarque 2

Dans les démonstrations, sans perte de généralité, on considérera uniquement les parties de $X(\Omega)$ et non de $\mathbb{R}$ pour les rendre plus lisibles.

Exemple 3 (*Jeu*)

Par exemple, l'ensemble des issues donnant une somme à 7 est :

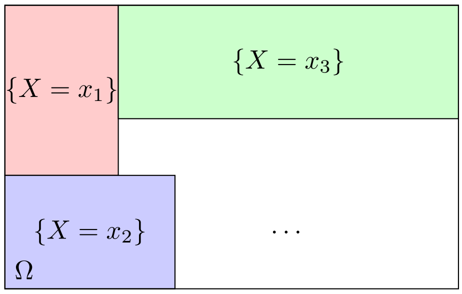
$$\{S = 7\} = S^{-1}(\{7\}) = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}.$$

et l'ensemble des issues donnant une somme à 2 ou 12 est :

$$\{S \in \{2, 12\}\} = S^{-1}(\{2, 12\}) = \{(1,1), (6,6)\}.$$

Proposition 1

Soit X une variable aléatoire.
L'ensemble $(X = x_i)_{1 \leqslant i \leqslant n}$ forme un système complet d'événements de Ω.



Démonstration :
- *disjoints :* Soit $i \neq j$. $\{i\}$ et $\{j\}$ sont deux ensembles disjoints. Donc $X^{-1}(\{i\})$ et $X^{-1}(\{j\})$ sont aussi disjoints.
- $\cup_{i=1}^{n}(X = x_i) = \cup_{i=1}^{n} X^{-1}(\{x_i\}) \overbrace{=}^{\text{Formule de Hausdorff}} X^{-1}(\cup_{i=1}^{n}\{x_i\}) = X^{-1}(X(\Omega)) = \Omega.$ ■

B Loi de probabilité

Définition-Proposition 1 (*Loi d'une variable aléatoire*)

Soit X une variable aléatoire.
L'application

$$P_X \left| \begin{array}{rcl} \mathcal{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & P(X^{-1}(A)) \end{array} \right.$$

est une probabilité sur $X(\Omega)$, appelée **loi de la variable X**.

Démonstration :
- $P_X(X(\Omega)) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1$
- Soit A et B deux événements incompatibles de $X(\Omega)$. Ainsi $X^{-1}(A)$ et $X^{-1}(B)$ sont deux événements incompatibles de Ω. On a

$$P_X(A \cup B) = P(X^{-1}(A \cup B)) = P(X^{-1}(A) \cup X^{-1}(B))$$

Comme $X^{-1}(A)$ et $X^{-1}(B)$ incompatibles, on a :

$$P_X(A \cup B) = P(X^{-1}(A)) + P(X^{-1}(B)) = P_X(A) + P_X(B).$$

■

Proposition 2 (*Déterminer une loi*)

*Déterminer **la loi de probabilité P_X** de la variable aléatoire X, c'est donner*

1. *l'ensemble $X(\Omega) = \{x_1, \ldots, x_n\}$ des valeurs prises par X,*
2. *pour chaque x_i de $X(\Omega)$, la probabilité $p_i = P(X = x_i)$.*

:

Démonstration : Soit $A \in \mathcal{P}(X(\Omega))$. Il existe $I \subset [\![1, n]\!]$ tel que $A = \{x_i : i \in I\}$.
On a

$$P(X \in A) = P(\bigcup_{i \in I}(X = x_i)) \overbrace{=}^{(\{X = x_i\})_{i \in I} \text{ disjoints}} \sum_{i \in I} P(X = x_i).$$ ■

Remarque 3

- $\forall i \in [\![1, n]\!] : \quad p_i = \sum_{\omega \in \Omega \text{ tel que } X(\omega) = x_i} P(\{\omega\}),$

— Comme les événements $X = x_1, X = x_2, \dots, X = x_n$ forme un système complet d'événements de Ω, on a :
$$\sum_{x_i \in X(\Omega)} P(X = x_i) = \sum_{i=1}^{n} p_i = 1.$$

Exemple 4 (*Jeu*)

La variable aléatoire somme $S \left| \begin{array}{rcl} \Omega & \longrightarrow & \{2, 3, \dots, 12\} \\ (b, r) & \longmapsto & b + r \end{array} \right.$ est déterminée par ce tableau :

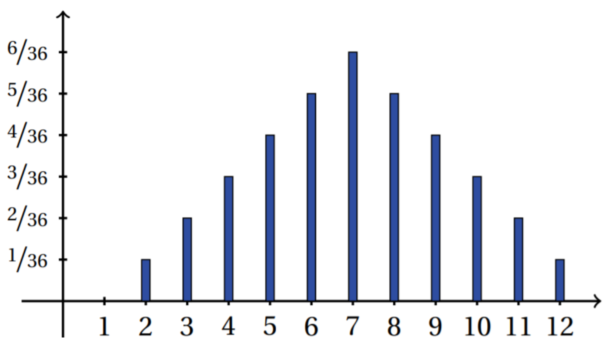
b \ r	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
2	5	6	7	8	9	10
1	6	7	8	9	10	11
2	7	8	9	10	11	12

Du fait de l'équiprobabilité, on détermine la loi de probabilité S en calculant le nombre de cases. Par exemple,
$$P(S = 4) = P(\{(1, 3), (2, 2), (3, 1)\}) = \frac{\text{card}(\{(1, 3), (2, 2), (3, 1)\})}{\text{card}(\{1, \dots, 6\}^2)} = \frac{3}{36}.$$

On obtient :

k	2	3	4	5	6	7	8	9	10	11	12
$p_k = P(S = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Sa représentation graphique est :



Théorème 3 (*Construire un espace de probabilité à partir d'une loi*)

Soit $(x_1, \dots, x_n)$ une famille finie de réels et $(p_1, \dots, p_n)$ une famille de réels positifs telle que $\sum_{i=1} p_i = 1$.
Alors il existe un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ et une variable aléatoire X sur cette espace et à valeurs dans $\{x_1, \dots, x_n\}$ telle que :
$$\forall i \in [\![1, n]\!] : \quad P(X = x_i) = p_i.$$

Démonstration : Soit $\Omega = \{x_1, \dots, x_n\}$ et P la probabilité définie sur $(\Omega, \mathcal{P}(\Omega))$ par :
$$\forall i \in [\![1, n]\!] : \quad P(\{x_i\}) = p_i.$$

Soit X l'application identité. On vérifie que $P(X = x_i) = P(\{x_i\}) = p_i$, pour tout $i \in [\![1, n]\!]$. ■

Remarque 4

Ce théorème permet de définir une variable aléatoire par sa loi de probabilité sans avoir à étudier l'expérience aléatoire sous-jacente, c'est à dire définir l'espace de probabilité $(\Omega, \mathcal{P}(\Omega), P)$ et la fonction de la variable aléatoire. Par exemple, on modélise l'expérience aléatoire d'un lancer de dé en posant : la variable aléatoire X représente le chiffre du dé avec $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$.

Définition-Proposition 2 (*Fonction d'une variable aléatoire*)

Soit X une variable aléatoire.
Soit $f : \mathbb{R} \to \mathbb{R}$ une application quelconque.
L'application $f \circ X$

$$f \circ X \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & f(X(\omega)) \end{array} \right.$$

est une variable aléatoire réelle.
L'usage veut qu'on la note abusivement $f(X)$ au lieu de $f \circ X$.

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[rr, bend right, "f \circ X"'] & \mathbb{R} \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

On a

$$\forall y \in f(X)(\Omega) : \quad P(f(X) = y) = \sum_{x \in X(\Omega) \text{ tel que } f(x) = y} P(X = x) = \sum_{i \in [\![1,n]\!] \text{ tel que } f(x_i) = y} P(X = x_i).$$

Exemple 5 (*Jeu*)

Le gain obtenu est fonction de la somme obtenue avec les deux dés. La modélisation est d'appliquer une fonction à la somme, s, des dés :

$$G \left| \begin{array}{ccc} [\![2, 12]\!] & \longrightarrow & \{-1, 0, 2\} \\ s & \longmapsto & \begin{cases} 2 & \text{si } s = 2 \text{ ou } 12 \\ -1 & \text{si } s = 7 \\ 0 & \text{sinon} \end{cases} \end{array} \right. .$$

La variable aléatoire Gain est $G \circ X$.
Vérifions sur un exemple la modélisation. Si le jet des dés donne $\omega = (1, 6)$, le gain de -1. On a bien $G \circ S(\omega = (1, 6)) = G(S(1, 6)) = G(1 + 6) = G(7) = -1$.
La loi de probabilité S est déterminé par :

k	-1	2	0
$P(G(S) = k)$	$\frac{6}{36} = P(S = 7)$	$\frac{2}{36} = P(S = 2) + P(S = 12)$	$\frac{28}{36} = 1 - P(G(S) = -1) + P(G(S) = 2)$

C Vecteurs aléatoires

Définition-Proposition 3 (*Loi conjointe*)

Soit X, Y deux variables aléatoires.

On note (X,Y) le **couple de variables aléatoires** prenant ses valeurs dans $\mathbb{R}^2$.
La **loi conjointe** du couple (X,Y) est déterminé par :

1. $X(\Omega)\times Y(\Omega)=\{x_1,\dots,x_n\}\times\{y_1,\dots,y_m\}$, les valeurs prises par le couple
2. $\forall(i,j)\in [\![1,n]\!]\times[\![1,m]\!]: \quad P(X=x_i\cap Y=y_j)$.

On note $p_{i,j}=P(X=x_i\cap Y=y_j)$.
Les événements $((X=x_i)\cap(Y=y_j))_{\forall(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ forment un système complet d'événements de Ω. En particulier, on a :

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

Démonstration : Démontrons que la loi conjointe est déterminée par

$$\forall(i,j)\in[\![1,n]\!]\times[\![1,m]\!]: \quad P(X=x_i\cap Y=y_j).$$

Soit $(A,B)\subset X(\Omega)\times Y(\Omega)$. Il existe $I\subset[\![1,n]\!]$ et $J\subset[\![1,m]\!]$ tel que $A=\{x_i : i\in I\}$ et $B=\{y_j : j\in J\}$.
On a :

$$P((X,Y)\in(A,B))=P(X\in A\cap Y\in B)=P\left(\left(\bigcup_{i\in I}(X=x_i)\right)\cap\left(\bigcup_{j\in J}(Y=y_j)\right)\right)=P(\bigcup_{i\in I,j\in J}(X=x_i)\cap(Y=y_j))$$

. Comme $(\{(X=x_i)\cap(Y=y_j)\})_{i\in I,j\in J}$ est une famille d'ensembles disjoints. On a

$$P((X,Y)\in(A,B))=P(X\in A\cap Y\in B)=\sum_{i\in I,j\in J}P(X=x_i\cap Y=y_j).$$

■

Définition 3 (*Lois marginales*)

Soit (X,Y) un couple de variables aléatoires.
La loi de X appelé **première loi marginale** et la loi de Y **second loi marginale** du couple.

Proposition 4 (*Relations*)

Soit (X,Y) un couple de variables aléatoires.
On a :

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

et

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

X \ Y	y_1	$\dots$	y_j	$\dots$	y_m	$P(X=x_i)$	
x_1	$p_{1,1}$	$\dots$	$p_{1,j}$	$\dots$	$p_{1,m}$	$P(X=x_1)$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	
x_i	$p_{i,1}$	$\dots$	$p_{i,j}$	$\dots$	$p_{i,m}$	$P(X=x_i)$	$\leftarrow \sum_{k=1}^{m}p_{i,k}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	
x_n	$p_{n,1}$	$\dots$	$p_{n,j}$	$\dots$	$p_{n,m}$	$P(X=x_n)$	
$P(Y=y_j)$	$P(Y=y_1)$	$\dots$	$P(Y=y_j)$	$\dots$	$P(Y=y_m)$	*1*	
			$\uparrow$ $\sum_{k=1}^{n}p_{k,j}$				

Démonstration : Comme les événements $((Y = y_j))_{\forall j \in [\![1,m]\!]}$ forment un système complet d'événements de Ω, d'après la proposition **??** on a

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

. ■

Exemple 6

On tire deux nombres au hasard dans $\{-1, 1\}$. On note X leur somme, et Y leur produit. On cherche à déterminer la loi conjointe de (X, Y).
L'espace de probabilité est est $(\{-1; 1\}^2, \mathcal{P}(\{-1; 1\}^2), P)$ avec P la probabilité uniforme. Pour tout $\omega = (r, b) \in \Omega$, on a $X(r, b) = r + b$ et $Y(r, b) = r.b$. Sous forme de tableau, on a :

r \ b	-1	1
-1	-2	0
1	0	2

Définition de X

r \ b	-1	1
-1	1	-1
1	-1	1

Définition de Y

X prend ses valeurs dans $\{-2, 0, 2\}$ et Y dans $\{-1, 1\}$.
Les variables aléatoires étant finies, il suffit de déterminer toutes les probabilités $P(X = x \text{ et } Y = y)$ pour tout couple $(x, y) \in \{-2, 0, 2\} \times \{-1, 1\}$. On a :

- $P(X = 2, Y = 1) = P(\{(1, 1)\}) \overbrace{=}^{\text{equiprobabilité}} \frac{1}{4}$.
- $P(X = 2, Y = -1) = P(\emptyset) = 0$.
- $P(X = 0, Y = 1) = P(\emptyset) = 0$.
- $P(X = 0, Y = -1) = P(\{(1, -1), (-1, 1)\}) = \frac{2}{4} = \frac{1}{2}$.
- $P(X = -2, Y = 1) = P(\{(-1, -1)\}) = \frac{1}{4}$.
- $P(X = -2, Y = -1) = P(\emptyset) = 0$.

La loi de probabilité conjointe (X, Y) est donc déterminée par :

X \ Y	-1	1	$P(X = x)$
-2	0	1/4	1/4
0	1/2	0	1/2
2	0	1/4	1/4
$P(Y = y)$	1/2	1/2	1

Exemple 7

Dans une classe, la répartition en fonction de l'age et du genre est :

Genre / Age	Fille	Garçon	Total
18	5	10	15
19	2	6	8
20	0	1	1
Total	7	17	24

On tire au hasard un élève dans la classe.
X représente l'age de l'élève.
Y représente le genre de l'élève avec la label 0 pour une fille et le label 1 pour un garçon.
La loi de probabilité conjointe (X, Y) est déterminée par :

$X \backslash Y$	0	1	$P(X = x)$
18	$5/24$	$10/24$	$15/24$
19	$2/24$	$6/24$	$8/24$
20	$0/24$	$1/24$	$1/24$
$P(Y = y)$	$7/24$	$17/24$	1

D Indépendance et lois conditionnelles

Définition-Proposition 4 (*Lois conditionnelles*)

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

1. On appelle **loi conditionnelle** de X sachant $(Y = y_j)$ pour $P(Y = y_j) \neq 0$, la probabilité définie par :
$$\forall i \in [\![1, n]\!] : \quad P_{Y=y_j}(X = x_i) = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_i)}$$
2. On appelle **loi conditionnelle** de Y sachant $(X = x_i)$ pour $P(X = x_i) \neq 0$, la probabilité définie par :
$$\forall j \in [\![1, m]\!] : \quad P_{X=x_i}(Y = y_j) = \frac{P(X = x_i \cap Y = y_j)}{P(X = x_i)}$$

Remarque 5

Les lois marginales et les lois conditionnelles déterminent la loi conjointe :
$$\forall (i, j) \in [\![1, n]\!] \times [\![1, m]\!] : \quad P(X = x_i \cap Y = y_j) = P_{X=x_i}(Y = y_j)P(X = x_i)$$
$$\forall (i, j) \in [\![1, n]\!] \times [\![1, m]\!] : \quad P(X = x_i \cap Y = y_j) = P_{Y=y_j}(X = x_i)P(Y = y_j)$$

Définition 4 (*Indépendance*)

Soit (X, Y) un couple de variables aléatoires.
On dit que X et Y sont **indépendantes** si
$$\forall (i, j) \in [\![1, n]\!] \times [\![1, m]\!] : \quad P((X = x_i) \cap (Y = y_j)) = P(X = x_i)P(Y = y_j).$$

Remarque 6

Dans la modélisation d'une expérience aléatoire, l'hypothèse d'indépendance des variables aléatoires est souvent une donnée de l'expérience et non pas une propriété à vérifier. Par exemple, la modélisation de l'expérience du jeter de deux dés serait :
- X_1, la variable aléatoire représentant le chiffre du premier dé,
- X_2, la variable aléatoire représentant le chiffre du second dé,
- X_1 et X_2, supposées indépendantes,

— $S = X_1 + X_2$, la somme des deux dés.

Théorème 5

Soit (X, Y) un couple de variables aléatoires.
Si X et Y sont indépendantes, il en va de même pour les variables aléatoires $f(X)$ et $g(Y)$ où $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ sont deux applications quelconques.

E Espérance

L'espérance mathématique d'une variable aléatoire réelle est la moyenne pondérée par les probabilités d'apparition de chaque valeur. Le théorème de la loi forte des grands nombres démontrera que l'espérance est la valeur que l'on s'attend à trouver, en moyenne, si l'on répète un grand nombre de fois la même expérience aléatoire.

Définition 5 (*Espérance d'une variable aléatoire*)

Soit X une variable aléatoire.
L'**espérance** de X est

$$E(X) = \sum_{\forall i \in [\![1,n]\!]} x_i P(X = x_i)$$

Exemple 8 (*Somme de deux dés*)

Reprenons l'exemple du lancer de deux dés en notant S la somme des chiffres obtenus. L'espérance de S vaut alors :

$$E(S) = \sum_{s \in S(\Omega)} sP(S = s) = 2.\frac{1}{36}+3.\frac{2}{36}+4.\frac{3}{36}+5.\frac{4}{36}+6.\frac{5}{36}+7.\frac{6}{36}+8.\frac{5}{36}+9.\frac{4}{36}+10.\frac{3}{36}+11.\frac{2}{36}+12.\frac{1}{36} = 7.$$

Théorème 6 (*Formule de transfert*)

Soit X une variable aléatoire et $f : \mathbb{R} \to \mathbb{R}$.
On a

$$E(f(X)) = \sum_{\forall i \in [\![1,n]\!]} f(x_i)P(X = x_i).$$

Démonstration :

$$E(f(X)) = \sum_{\forall j \in [\![1,m]\!]} y_j P(f(X) = y_i) = \sum_{\forall j \in [\![1,m]\!]} y_j \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} P(X = x_i) = \sum_{\forall j \in [\![1,m]\!]} \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} y_j P(X = x_i).$$

D'où

$$E(f(X)) = \sum_{\forall j \in [\![1,m]\!]} \sum_{\forall i \in [\![1,n]\!] : f(x_i) = y_j} f(x_i)P(X = x_i) = \sum_{\forall i \in [\![1,n]\!]} f(x_i)P(X = x_i).$$

■

Remarque 7

La formule de transfert permet de calculer $E(f(X))$ sans avoir à déterminer la loi de $f(X)$.
Dans l'exemple du jeu, l'espérance de la variable aléatoire $G(X)$ représente le gain moyen du joueur. On applique la formule du transfert :

$$E(G(S)) = \sum_{s \in \{2,3,\dots,12\}} G(s)P(S = s) = 2P(S = 2) - 1P(S = 7) + 2P(S = 12) = 2\frac{1}{36} - 1\frac{6}{36} + 2\frac{2}{36} = -\frac{1}{18}.$$

En conclusion, ce jeu n'est pas favorable au joueur.

Théorème 7

Soit X et Y deux variables aléatoires indépendantes.
Alors
$$E(XY) = E(X)E(Y).$$

Démonstration : On applique le théorème de transfert à la variable aléatoire $f(X,Y) = XY$ d'où

$$\begin{aligned}
E(XY) &= \sum_{\forall (i,j)\in[\![1,n]\!]\times[\![1,m]\!]} x_i y_j P(X = x_i \cap Y = y_j) \\
&\overbrace{=}^{\text{indépendance}} \sum_{\forall (i,j)\in[\![1,n]\!]\times[\![1,m]\!]} x_i y_j P(X = x_i) P(Y = y_j) \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} x_i y_j P(X = x_i) P(Y = y_j) \\
&= \sum_{i=0}^{n} x_i P(X = x_i) \sum_{j=0}^{m} y_j P(Y = y_j) \\
&= E(X)E(Y).
\end{aligned}$$

■

Proposition 8 (*propriétés de l'espérance*)

Soit X et Y deux variables aléatoires réelles et $(a,b) \in \mathbb{R}^2$.
- $E(a) = a$.
- *linéarité :* $E(aX + bY) = aE(X) + bE(Y)$.
- *Si* $P(X \geqslant 0) = 1$, *alors* $E(X) \geqslant 0$.
- *Si* $P(X \leqslant Y) = 1$, *alors* $E(X) \leqslant E(Y)$.

Définition 6 (*Centrée*)

Une variable aléatoire est dite **centrée** si son espérance est nulle.

Exemple 9 (*Centrer une variable aléatoire*)

Soit X une variable aléatoire.
La variable aléatoire $Y = X - E(X)$ est centrée car $E(Y) = E(X - E(X)) = E(X) - E(E(X)) = E(X) - E(X) = 0$

Démonstration :
- Dans l'expression $E(a)$, a est la variable aléatoire définie par $\omega \mapsto a$ d'où $E(a) = \overbrace{a}^{\in\mathbb{R}} P(a = \overbrace{a}^{\in\mathbb{R}}) = \overbrace{a}^{\in\mathbb{R}}$.
- On applique le théorème de transfert à la variable aléatoire $f(X,Y) = aX + bY$ d'où

$$\begin{aligned}
E(aX + bY) &= \sum_{\forall i\in[\![1,n]\!]} \sum_{\forall j\in[\![1,m]\!]} (ax_i + by_j) P(X = x_i \cap Y = y_j) \\
&= a \sum_{\forall i\in[\![1,n]\!]} x_i \sum_{\forall j\in[\![1,m]\!]} P(X = x_i \cap Y = y_j) + b \sum_{\forall j\in[\![1,m]\!]} y_j \sum_{\forall i\in[\![1,n]\!]} P(X = x_i \cap Y = y_j) \\
&= a \sum_{\forall i\in[\![1,n]\!]} x_i P(X = x_i) + b \sum_{\forall j\in[\![1,m]\!]} y_j P(Y = y_j) \\
&= aE(X) + bE(Y).
\end{aligned}$$

- Soit X une variable aléatoire telle que $P(X \geqslant 0) = 1$ d'où $P(X < 0) = 0$.
  De plus, pour tout $x \in \mathbb{R}$, $P(X = x) = 0$ car $\{X = x\} \subset \{X < 0\}$.

$$E(X) = \sum_{\forall i\in[\![1,n]\!]} x_i P(X = x_i) = \sum_{\forall i\in[\![1,n]\!] : x_i \geqslant 0} x_i P(X = x_i) \geqslant 0.$$

— Soit X une variable aléatoire telle que $P(Y-X \leqslant 0)=1$. Ainsi $E(Y-X) \geqslant 0$ et $E(X) \leqslant E(Y)$. ■

F Variance

La variance est une mesure de la dispersion des valeurs d'une loi de probabilité.

Définition 7 (*variance d'une variable aléatoire*)

Soit X une variable aléatoire.
La **variance** de X est

$$E\left((X-E(X))^2\right)$$

et on la note $V(X)$.
L'**écart-type** est $\sigma(X)=\sqrt{V(X)}$.

Proposition 9 (*Formule de Kœnig-Hugens*)

Soit X une variable aléatoire.
Pour tout $a>0$, on a :

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

Démonstration :

$$V(X)=E\left((X-E(X))^2\right)$$
$$V(X)=E\left(X^2-2E(X)X+E(X)^2\right)$$
$$V(X)=E(X^2)-2E(X)E(X)+E(X)^2 \text{ car } E \text{ linéaire et } E(a)=a$$
$$V(X)=E(X^2)-(E(X))^2$$

■

Exemple 10 (*Somme de deux dés*)

Reprenons l'exemple du lancer de deux dés en notant S la somme des chiffres obtenus. On a :

$$E(S^2) \overbrace{=}^{\text{th transfert}} \sum_{s\in S(\Omega)} s^2 P(S=s) = 2^2.\frac{1}{36}+3^2.\frac{2}{36}+4^2.\frac{3}{36}+\cdots+12^2.\frac{1}{36}=\frac{329}{6}.$$

La variance de S vaut alors :

$$V(S)=E(S^2)-(E(S))^2=\frac{329}{6}-7^2=\frac{35}{6}.$$

et son écart-type

$$\sigma(S)=\sqrt{V(S)}\approx 2,42.$$

Définition 8 (*Réduite*)

Une variable aléatoire est dite **réduite** si sa variance est égale à .

Exemple 11 (*Centrer et réduire une variable aléatoire*)

Soit X une variable aléatoire.
La variable aléatoire $Y=\frac{X-E(X)}{\sigma(X)}$ est centrée réduite.

Définition 9 (*Covariance de deux variables aléatoires*)

Soit X et Y deux variables aléatoires.

La **covariance** de X et Y est

$$E((X - E(X))(Y - E(Y)))$$

et on la note $\text{Cov}(X, Y)$.

Proposition 10 (*Propriétés*)

Soit X et Y deux variables aléatoires.

— $V(aX + b) = E\left((X - E(X))^2\right)$.

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

— *Si X et Y sont indépendantes, alors* $V(X + Y) = V(X) + V(Y)$.

Démonstration : — $V(aX + b) = E((aX + b - E(aX + b))^2) = E(a^2(X - E(X))^2) = a^2V(X)$ car E linéaire..

— $V(X + Y) = E((X + Y - E(X + Y))^2) = E((X - E(X))^2) + E((Y - E(Y))^2) + 2E((X - E(X)(Y - E(Y))) = Var(X) + Var(Y) + 2\text{Cov}(X, Y)$.

— Si X et Y sont indépendantes, alors $X - E(X)$ et $Y - E(Y)$ sont indépendantes, d'où

$$E((X - E(X)(Y - E(Y))) = E(X - E(X))E(Y - E(Y)) = 0.$$

■

Définition 10 (*Corrélation*)

Soit X et Y deux variables aléatoires de variance non nulles.
Le **coefficient de corrélation** de X et Y est

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

Proposition 11 (*Propriétés*)

Soit X et Y deux variables aléatoires de variance non nulles.

— $\rho(X, Y) \in [-1, 1]$

— $\rho(X, Y) = \pm 1$ *si et seulement si il existe deux réels a et b tels que l'événement* $(Y = aX + b)$ *soit certain.*

— *Si X et Y sont indépendantes, alors $\rho(X, Y) = 0$. On dit qu'elles sont décorrélées. La réciproque est fausse.*

G Lois usuelles

Définition 11 (*Loi uniforme*)

Une variable aléatoire X suit la **loi uniforme** sur $[\![1, n]\!]$ si

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

On note $X \hookrightarrow \mathcal{U}(E)$.
On a : $E(X) = \frac{n+1}{2}$ et $V(X) = \frac{n^2-1}{2}$.

Définition 12 (*Loi de Bernoulli*)

Une variable aléatoire X suit la **loi de Bernoulli** de paramètre p si

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

On note $X \hookrightarrow \mathcal{B}(p)$.
On a : $E(X) = p$ et $V(X) = p(1 - p)$.

Proposition 12 (*Loi indicatrice*)

Soit A est un événement de Ω. La variable aléatoire réelle 1_A définie par

$$1_A(\omega) = \begin{cases} 1 & si\ \omega \in A \\ 0 & si\ \omega \in \bar{A} \end{cases}$$

est une variable de Bernoulli de paramètre $p = P(A)$. On l'appelle ***loi indicatrice*** *de l'événement A.*

Définition 13 (*Loi binomiale*)

Une variable aléatoire X suit la **loi binomiale** de paramètres n et p si

1. $X(\Omega) = [\![1, n]\!]$,
2. $\forall k \in [\![1, n]\!]$: $\quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

On note $X \hookrightarrow \mathcal{B}(n, p)$.
On a : $E(X) = np$ et $V(X) = np(1-p)$.

Remarque 8

Pour $n = 1$, on retrouve la loi de Bernouilli de paramètre p.

Proposition 13 (*Loi des tirages avec remise*)

Soit $X_1, X_2, \ldots, X_n$ n variables aléatoires indépendantes de loi de Bernoulli de même paramètre p.
Alors $S = X_1 + X_2 + \cdots + X_n$ suit une loi binomiale de paramètres n, p.
En d'autres termes, si l'expérience aléatoire est une répétition de n épreuves identiques et indépendantes tel que chaque épreuve est une expérience Bernouilli de paramètre p et si S est égale au nombre de succès des n épreuves de l'expérience, alors S suit une loi binomiale de paramètres n, p.

H Convergence et approximations

Proposition 14 (*Inégalité de Markov*)

Soit X une variable aléatoire réelle positive.
Pour tout $a > 0$, on a :

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

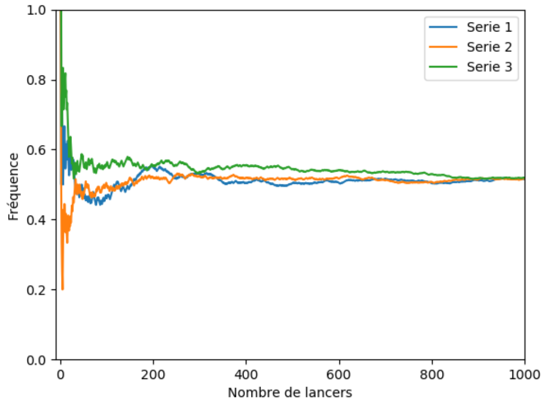
Proposition 15 (*Inégalité de Bienaymé-Tchebychev*)

Soit X une variable aléatoire réelle. Pour tout $a > 0$, on a

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

Cette inégalité présente un intérêt théorique en majorant la probabilité qu'une variable aléatoire s'écarte de sa moyenne. Nous allons l'utiliser pour prouver la loi faible des grands nombres.

Lors d'un lancer d'une pièce de monnaie équilibrée, les deux côtés « pile » et « face » apparaissent de façon équiprobable pour des raisons de symétrie : on ne s'attend pas plus à l'un ou à l'autre côté. Cette mesure de l'attente s'appuie souvent sur une considération statistique : on observe que la fréquence des occurrences de chaque côté se rapproche de $1/2$.



Le théorème de la loi des grands nombres permet de justifier ce résultat en interprétant la probabilité comme une fréquence de réalisation.
La modélisation de cette expérience aléatoire est :

1. **n$^{\text{ième}}$ lancer** : pour tout $n \in \mathbb{N}^*$, la variable aléatoire X_n, représente le résultat du n$^{\text{ième}}$ lancer. Elle suit une loi Bernouilli de paramètre p (l'issue 0 représente le pile et l'issue 1 le face). Elles sont supposées indépendantes. On a $E(X_n) = (1-p).0 + p.1 = p$.
2. **fréquence** : pour tout $n \in \mathbb{N}^*$, la variable aléatoire S_n, représente la moyenne des résultats obtenus au cours des n premiers lancers, soit :

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

L'issue, ω, correspondant à une succession de faces existe, $X_n(\omega) = 1$ pour tout $n \in \mathbb{N}^*$. Donc 1 est une valeur possible de la variable aléatoire S_n. Cependant, le théorème de la loi faible des grands nombres prouve que la probabilité que S_n s'écarte de l'espérance $E(X_n)$ tend vers 0 quand n tend vers l'infini.

Théorème 16 (*Loi faible des grands nombres*)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. On suppose toutes les variables indépendantes et de même loi, admettant une espérance m et un écart type σ. Posons $S_n = \frac{X_1+X_2+\cdots+X_n}{n}$.
Alors

$$\forall \epsilon > 0 : \quad P\left(|S_n - m| \geqslant \epsilon\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Démonstration : D'après l'inégalité de Bienaymé-Tchebychev et par indépendance des variables aléatoires,

$$\forall \epsilon > 0 : \quad P\left(|S_n - m| \geqslant \epsilon\right) \leqslant \frac{V(S_n)}{\epsilon^2} = \frac{\frac{V(X_1+X_2+\cdots+X_n)}{n^2}}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \underset{n\to+\infty}{\longrightarrow} 0.$$ ■

Remarque 9

Le TP Python `https://github.com/VincentTariel/cours/blob/master/probabilite/simulation_variable_aleatoire_avtivite_python.pdf` permet de vous familiariser avec ce théorème sur des applications.