

Chapitre 4 - Probabilités - Variables aléatoires discrètes

I – Pour pouvoir sommer tranquillement ...

Les résultats de cette section seront admis pour la plupart. L'objectif de cette section est uniquement de fournir un cadre et un vocabulaire rigoureux qui permettent d'étendre les résultats de première année sur les sommes finies.

En effet, en deuxième année, l'univers Ω n'est plus supposé fini ...

1) Ensembles au plus dénombrables

Définition I.1.1. • Un ensemble X est dit dénombrable s'il existe une bijection de $\mathbb{N}$ dans X , autrement dit si on peut écrire $X = \{x_n, n \in \mathbb{N}\}$ où les x_n sont distincts.

- Un ensemble X est dit au plus dénombrable s'il est fini ou dénombrable.
- Une famille $(x_i)_{i \in I}$ indexée sur un ensemble I est dite (au plus) dénombrable si l'ensemble I des indices est (au plus) dénombrable.

Remarque. Un ensemble X est au plus dénombrable si et seulement s'il existe une bijection entre X et un ensemble de la forme $\llbracket 1; n \rrbracket$ (X fini) ou entre X et $\mathbb{N}$ (X dénombrable).

Dans les deux cas, cela permet de *numéroter* ses éléments, cette numérotation se faisant sur un ensemble fini ou pas.

Exemple I.1.2. (1) Un ensemble en bijection avec un ensemble dénombrable est dénombrable.

(2) Sont dénombrables : $\mathbb{N}, \mathbb{N}^*, 2\mathbb{N}$ et toute partie infinie de $\mathbb{N}; \mathbb{Z}; \mathbb{N} \times \mathbb{N}, \mathbb{Q}$.

(3) Tout sous-ensemble d'un ensemble dénombrable est au plus dénombrable.

Une réunion au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable.

Un produit cartésien d'un nombre fini d'ensembles au plus dénombrables est au plus dénombrable (comme $\mathbb{N} \times \mathbb{N}$).

L'image d'un ensemble dénombrable par une application est au plus dénombrable.

(4) Sont infinis et non dénombrables : $]0; 1[, \mathbb{R}, \{0, 1\}^{\mathbb{N}}$ (ensemble des suites à valeurs dans $\{0, 1\}$).

2) Familles dénombrables sommables

Dans cette partie, on considère des familles dénombrables de nombres réels ou complexes même si, dans la pratique, ces familles seront la plupart du temps des familles de nombres réels, souvent positifs.

Définition - Théorème I.2.3.

Soit $(x_i)_{i \in I}$ une famille **dénombrable** de nombres réels/complexes. Il existe donc une bijection $\varphi : \mathbb{N} \rightarrow I$.

Le choix de cette bijection permet de mettre dans *un certain ordre* les x_i en formant une suite $(x_{\varphi(n)})_{n \in \mathbb{N}}$ dont les termes sont les x_i .

On dit que la famille $(x_i)_{i \in I}$ est sommable si la série $\sum x_{\varphi(n)}$ est **absolument** convergente. La convergence absolue et la somme de cette série ne dépendent pas de la bijection φ choisie. On note alors simplement $\sum_{i \in I} x_i$ la somme de cette série, appelée aussi somme de la famille $(x_i)_{i \in I}$.

Dans la pratique, on peut ne pas expliciter la bijection φ et n'utiliser que cette dernière notation et les propriétés de la page suivante.

Proposition I.2.4 (Théorème de comparaison de familles dénombrables - ADMIS).

Soit $(x_i)_{i \in I}$ une famille dénombrable de nombres réels/complexes. Soit $(y_i)_{i \in I}$ une famille de réels positifs.

On suppose que, pour tout $i \in I$, $|x_i| \leq y_i$.

Si la famille $(y_i)_{i \in I}$ est sommable, alors la famille $(x_i)_{i \in I}$ est sommable.

Théorème I.5 (ADMIS). En cas de sommabilité, les sommes de familles dénombrables se manipulent naturellement comme le détaillent les propriétés suivantes :

- (1) (inégalité triangulaire) Soit $(x_i)_{i \in I}$ une famille dénombrable sommable de nombres réels/complexes.

$$\text{On a } \left| \sum_{i \in I} x_i \right| \leq \sum_{i \in I} |x_i|.$$

- (2) (croissance) Soient $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ deux familles dénombrables sommables de nombres réels.

On suppose que $x_i \leq y_i$ pour tout $i \in I$. Alors $\sum_{i \in I} x_i \leq \sum_{i \in I} y_i$.

- (3) (linéarité) Soient $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ deux familles dénombrables sommables de nombres réels/complexes. Soit $\lambda \in \mathbb{C}$.

Alors la famille $(x_i + \lambda y_i)_{i \in I}$ est sommable et $\sum_{i \in I} (x_i + \lambda y_i) = \left(\sum_{i \in I} x_i \right) + \lambda \left(\sum_{i \in I} y_i \right)$.

Théorème I.6 (Sommutation par paquets - ADMIS).

- (1) Soit I un ensemble dénombrable. Soit $(I_k)_{1 \leq k \leq n}$ une partition de I (chaque élément de I appartient à un et un seul I_k).

Une famille de nombres réels/complexes $(x_i)_{i \in I}$ est sommable si et seulement si les familles $(x_i)_{i \in I_k}$ sont finies

ou sommables. On a alors $\sum_{i \in I} x_i = \sum_{k=1}^n \sum_{i \in I_k} x_i$. On parle alors de sommutation par paquets (un nombre fini de paquets ici).

- (2) Soit I un ensemble dénombrable. Soit $(I_n)_{n \in \mathbb{N}}$ une partition de I (chaque élément de I appartient à un et un seul I_n).

Une famille de nombres réels/complexes $(x_i)_{i \in I}$ est sommable si et seulement si :

- pour tout $n \in \mathbb{N}$, les familles $(x_i)_{i \in I_n}$ sont finies ou sommables;

- la famille $\left(\sum_{i \in I_n} |x_i| \right)_{n \in \mathbb{N}}$ est une famille sommable.

On a alors $\sum_{i \in I} x_i = \sum_{n=0}^{+\infty} \sum_{i \in I_n} x_i$. On parle encore de sommutation par paquets (une infinité dénombrable de paquets ici).

Remarque.  On reconnaît dans les propriétés des familles sommables les propriétés couramment utilisées en première année pour des sommes finies : invariance de la somme par permutation des termes, linéarité, sommutation par paquets (= commutativité + associativité!)

On rappelle que l'hypothèse fondamentale pour pouvoir utiliser ces propriétés est **l'absolue convergence**.

Bien sûr, dans le cas de **familles finies** (probabilités de première année), toutes ces propriétés sont immédiates (commutativité, associativité de l'addition; distributivité de la multiplication sur l'addition) et continueront à être utilisées sans aucune justification!

 Dans le cas de familles dénombrables de **réels positifs**, on pourra **directement** entreprendre les calculs ou les majorations en utilisant ces propriétés, le fait que la somme soit finalement finie suffisant à justifier *a posteriori* la sommabilité.

Dans le cas contraire (famille de **réels positifs** non sommable), le calcul fera apparaître une série divergente, c'est-à-dire une somme égale à $+\infty$.

Méthode.  Pour montrer la sommabilité et calculer la somme d'une famille dénombrable $(x_i)_{i \in I}$:

- on utilise le cours sur les séries si $I = \mathbb{N}$ ou $I = \mathbb{N}^* \dots$;
- on utilise la sommation par paquets si $I = \mathbb{Z}$, $I = \mathbb{N} \times \mathbb{N} \dots$

Si les termes sont positifs (souvent!), on peut entreprendre directement le calcul de la somme, la finitude du résultat valant preuve de sommabilité.

Dans le cas général, on commence par montrer que $(|x_i|)_{i \in I}$ est sommable (en sommant et trouvant un résultat fini) puis on calcule la somme.

- on utilise une comparaison entre familles dénombrables :
si $|x_i| \leq y_i$ pour tout $i \in I$, la sommabilité de $(y_i)_{i \in I}$ entraîne celle de $(x_i)_{i \in I}$ mais ne donne pas la somme ...

Exercice I.2.7. (1) Montrer que la famille $\left(\frac{1}{n^2}\right)_{n \in \mathbb{Z}^*}$ est sommable et calculer sa somme.

(2) Soit $\alpha \in \mathbb{R}$. La famille $\left(\frac{(-1)^{p+q}}{(p+q+1)^\alpha}\right)_{(p,q) \in \mathbb{N}^2}$ est-elle sommable si $\alpha = 2$? si $\alpha = 3$?

Théorème I.8 (Fubini - Sommes doubles). Soient I et J deux ensembles au plus dénombrables.

Soit $(u_{i,j})_{(i,j) \in I \times J}$ une famille (au plus dénombrable donc) de nombres réels/complexes.

(1) La famille $(u_{i,j})_{(i,j) \in I \times J}$ est finie ou sommable si et seulement si :

- pour tout $j \in J$, la famille $(u_{i,j})_{i \in I}$ est finie ou sommable;
- la famille $\left(\sum_{i \in I} |u_{i,j}|\right)_{j \in J}$ est finie ou sommable.

$$\text{On a alors } \sum_{(i,j) \in I \times J} u_{i,j} = \sum_{j \in J} \sum_{i \in I} u_{i,j}.$$

(2) La famille $(u_{i,j})_{(i,j) \in I \times J}$ est finie ou sommable si et seulement si :

- pour tout $i \in I$, la famille $(u_{i,j})_{j \in J}$ est finie ou sommable;
- la famille $\left(\sum_{j \in J} |u_{i,j}|\right)_{i \in I}$ est finie ou sommable.

$$\text{On a alors } \sum_{(i,j) \in I \times J} u_{i,j} = \sum_{i \in I} \sum_{j \in J} u_{i,j}.$$

Dans le cas d'une famille $(u_{i,j})_{(i,j) \in I \times J}$ finie ou sommable, on peut échanger les symboles $\sum$.

Exercice I.2.9. (1) Montrer que la famille $\left(\frac{(-3)^{p+q}}{p!q!}\right)_{(p,q) \in \mathbb{N}^2}$ est sommable et calculer sa somme.

(2) Démontrer l'existence et calculer $\sum_{(p,q) \in \mathbb{N} \times \mathbb{N}^*} \frac{1}{(p+q^2)(p+q^2+1)}.$

Corollaire I.2.10 (Produit de sommes). Soient I et J deux ensembles au plus dénombrables. Soient $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ deux familles de nombres réels/complexes.

Si les familles $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ sont finies ou sommables, alors la famille $(x_i y_j)_{(i,j) \in I \times J}$ est finie ou sommable

$$\text{et on a } \sum_{(i,j) \in I \times J} x_i y_j = \left(\sum_{i \in I} x_i \right) \left(\sum_{j \in J} y_j \right).$$

Corollaire I.2.11 (Produit de Cauchy de séries absolument convergentes). Soit $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries numériques.

$$\text{Pour tout } n \in \mathbb{N}, \text{ on note } w_n = \sum_{p+q=n} u_p v_q = \sum_{p=0}^n u_p v_{n-p}.$$

Si les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ convergent absolument, alors leur produit de Cauchy $\sum w_n$ converge absolument et :

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n \right) \times \left(\sum_{n=0}^{+\infty} v_n \right).$$

II – Univers des issues d’une expérience aléatoire

Dans ce chapitre, les commentaires *en italique* ne font pas partie intégrante du cours mais aident à sa compréhension, en particulier en faisant le lien entre une approche intuitive des probabilités et le formalisme de la théorie des probabilités.

1) Expérience aléatoire et univers

Définition II.1.1. On appelle expérience aléatoire une expérience renouvelable et qui, renouvelée dans des conditions identiques (en principe sinon en pratique), ne donne pas forcément le même résultat.

Exemple II.1.2. (1) lancer un dé, deux dés,...

- (2) lancer indéfiniment une pièce et noter les résultats obtenus (Pile ou Face),
- (3) connaître le nombre de buts marqués lors un match de football,
- (4) attendre un bus à l’arrêt Meunier à partir de 17h ...

Remarquons que, quand nous parlons de **conditions identiques**, il faut comprendre : **identiques, pour autant que l’observateur puisse s’en assurer**. En général, les expériences réalisées sont parfaitement **déterministes**. Dans le cadre de la mécanique du solide par exemple, si la position initiale du dé, le support sur lequel il tombe et les conditions météorologiques sont identiques, le dé devrait tomber sur la même face! Il y a dans la pratique tant d’imprécisions dans la connaissance de ces différents paramètres que cela interdit la prévision du résultat et justifie l’emploi des expressions **aléatoire** ou **au hasard**.

Définition II.1.3. On appelle univers et on note traditionnellement Ω l’ensemble des résultats possibles (on dit des *issues*) d’une expérience aléatoire.

Toutes les expériences aléatoires se traduisent donc par le choix “au hasard” d’un élément ω dans un ensemble Ω . Dans la suite, Ω sera toujours supposé **non vide**. L’objectif de la suite du cours va être de préciser les modalités de ce choix *au hasard* ...

Exemple II.1.4. • Dans le cas du lancer d’un de dé à 6 faces, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- Dans le cas du lancer de 2 dés à 6 faces, $\Omega = \{1, 2, 3, 4, 5, 6\}^2$.
- Dans le cas du lancer d’une pièce indéfiniment, $\Omega = \{P, F\}^{\mathbb{N}}$ (ensembles des suites dont les termes sont des P et des F).
- Dans le cas du temps d’attente à un arrêt de bus, $\Omega = \mathbb{R}_+$.
- Dans le cas du nombre de buts, $\Omega = \mathbb{N}$ (on pourrait choisir $\mathbb{N} \times \mathbb{N}$ si on observait le score).

2) Vocabulaire des événements

Certains faits liés au résultat de l'expérience aléatoire peuvent se produire ou non. On les appellera des **événements**. Ils sont en général décrits par des assertions (des phrases), vraies pour certains $\omega \in \Omega$ et fausses pour les autres. Par exemple, dans le cas d'un jeu de Pile ou Face infini, l'événement A : "le premier Pile est obtenu avant le 10^e tirage". Les issues pour lesquelles l'événement A est réalisé forment un sous-ensemble de Ω qui représente l'événement et que, traditionnellement, on note encore A . Dans l'exemple précédent (comme souvent), il est difficile d'expliciter ce sous-ensemble : A est l'ensemble des suites de P et de F pour lesquelles l'un des 10 premiers termes est un P !

Définition II.2.5. Soit une expérience aléatoire associée à un univers Ω non vide.

- On appelle tribu sur Ω toute famille $\mathcal{A}$ de parties de Ω qui satisfait les propriétés suivantes :
 - $\Omega \in \mathcal{A}$;
 - si $A \in \mathcal{A}$, alors $\bar{A} \in \mathcal{A}$ (on dit que $\mathcal{A}$ est stable par passage au complémentaire) ;
 - pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{A}$, on a $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ (on dit que $\mathcal{A}$ est stable par réunion dénombrable).
- On appelle alors espace probabilisable un couple $(\Omega, \mathcal{A})$ où $\mathcal{A}$ est une tribu sur Ω . Les éléments de la tribu $\mathcal{A}$ sont alors appelés événements.
- Souvent, on appelle événement (et on note encore A) une assertion qui caractérise les éléments de A , c'est-à-dire qui est réalisée pour les issues qui sont des éléments de A et non réalisée pour les autres éléments de Ω . Soit $\omega \in \Omega$. On dit alors que l'événement A est réalisé par ω (A est alors une assertion!) si et seulement si $\omega \in A$ (A est alors un sous-ensemble de Ω !).
- On appelle événement élémentaire un événement de la forme $\{\omega\}$ où $\omega \in \Omega$ (un événement élémentaire est réalisé par une seule issue de l'expérience, c'est un *singleton*).

Remarque. L'ensemble $\mathcal{P}(\Omega)$ des parties de Ω est une tribu sur Ω . Dans le cas où Ω est fini (probabilités de première année), c'est toujours celle-là qui est choisie et il n'est même pas nécessaire de le préciser. Tous les sous-ensembles de Ω sont alors des événements! On pourrait faire de même lorsque Ω est dénombrable.

Dans le cas où Ω n'est pas dénombrable, des problèmes subtils et très largement hors-programme se posent et obligent à ne pas choisir $\mathcal{P}(\Omega)$ comme tribu. Ces questions ne seront pas soulevées en PSI et

la tribu $\mathcal{A}$ des événements n'aura jamais à être explicitée. En général, l'univers Ω ne le sera pas non plus!

Proposition II.2.6. Soit Ω un ensemble non vide. Soit $\mathcal{A}$ une tribu sur Ω .

- $\emptyset \in \mathcal{A}$;
- pour toute famille finie $(A_i)_{1 \leq i \leq n}$ d'éléments de $\mathcal{A}$, on a $\bigcup_{i=1}^n A_i \in \mathcal{A}$ (on dit que $\mathcal{A}$ est stable par réunion finie) ;
- pour toute famille finie $(A_i)_{1 \leq i \leq n}$ d'éléments de $\mathcal{A}$, on a $\bigcap_{i=1}^n A_i \in \mathcal{A}$ (on dit que $\mathcal{A}$ est stable par intersection finie) ;
- pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{A}$, on a $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$ (on dit que $\mathcal{A}$ est stable par intersection dénombrable) ;
- pour tous $A \in \mathcal{A}$ et $B \in \mathcal{A}$, on a $A \setminus B \in \mathcal{A}$.

Dans la définition suivante, on donne des correspondances entre les vocabulaires utilisés lorsqu'on considère les événements comme des sous-ensembles (voir la proposition précédente) ou comme des assertions :

Définition II.2.7. Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Soient A et B deux événements. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.

- Soit $\omega \in \Omega$. L'événement $A \cup B$ est réalisé (par ω) si et seulement si $\omega \in A$ ou $\omega \in B$, autrement dit si A est réalisé ou B est réalisé.
- Soit $\omega \in \Omega$. L'événement $\bigcup_{n \in \mathbb{N}} A_n$ est réalisé (par ω) si et seulement s'il existe $n \in \mathbb{N}$ tel que $\omega \in A_n$, autrement dit s'il existe $n \in \mathbb{N}$ tel que A_n est réalisé.
- Soit $\omega \in \Omega$. L'événement $A \cap B$ est réalisé (par ω) si et seulement si $\omega \in A$ et $\omega \in B$, autrement dit si A est réalisé et B est réalisé.
- Soit $\omega \in \Omega$. L'événement $\bigcap_{n \in \mathbb{N}} A_n$ est réalisé (par ω) si et seulement si $\omega \in A_n$ pour tout $n \in \mathbb{N}$, autrement dit si A_n est réalisé pour tout $n \in \mathbb{N}$.
- On appelle contraire de A l'événement $\bar{A}$ (le complémentaire de A dans Ω). Soit $\omega \in \Omega$. On dit que $\bar{A}$ est réalisé (par ω) si seulement si A n'est pas réalisé (par ω).
- Le fait d'avoir $A \subset B$ signifie que, si A est réalisé, alors B est réalisé (en termes d'assertions, cela signifie que A implique B).
- $\emptyset$ est appelé événement impossible; Ω est appelé événement certain. On dit que A et B sont incompatibles si $A \cap B = \emptyset$.

Exemple II.2.8. Considérons l'expérience aléatoire qui consiste à lancer une pièce indéfiniment.

Pour tout $n \in \mathbb{N}^*$, notons les événements

A_n : "obtenir Pile lors du n -ième lancer" et B_n : "obtenir Pile lors d'au moins un des n premiers lancers".

Notons également T : "obtenir au moins un Pile".

Si le jeu consiste à lancer la pièce jusqu'à obtenir un Pile, réaliser l'événement T signifie que le jeu se termine.

Exprimer T et $\bar{T}$ en fonction des événements A_n . Mêmes questions en fonction des événements B_n .

Quelles inclusions peut-on écrire entre les événements B_n ?

3) Vocabulaire des variables aléatoires

Lors d'une expérience aléatoire, le résultat peut être de nature très variée : nombre, couleur, lettre, individu, paire de chaussettes... L'utilisation d'une variable aléatoire réelle va permettre d'associer à chaque issue de l'expérience aléatoire un nombre réel, ce qui va permettre de pouvoir comparer les différents résultats de l'expérience, de faire des calculs sur ces résultats (espérance, variance,...), mais aussi de pouvoir modéliser de nombreuses situations a priori très différentes sous forme d'une des variables aléatoires "usuelles".

Définition II.3.9. Soit $(\Omega, \mathcal{A})$ un espace probabilisable.

Une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$ est une application définie sur Ω telle que $X(\Omega)$ soit au plus dénombrable et que, pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{A}$.

Cette dernière propriété permettra d'affirmer que $(X = x)$ est un événement !

L'ensemble $X(\Omega)$ des valeurs prises par X est appelé image de X . Lorsque $X(\Omega)$ est une partie de $\mathbb{R}$, on dit que la variable aléatoire est réelle.

Définition - Théorème II.3.10 (Notations).

Soit $(\Omega, \mathcal{A})$ un espace probabilisable.

Soient $X : \Omega \longrightarrow X(\Omega)$ une variable aléatoire discrète (réelle ou pas) et A une partie de $X(\Omega)$.

$X^{-1}(A)$ est un événement que l'on notera $\{X \in A\}$ ou $(X \in A)$.

C'est le sous-ensemble de Ω formé des issues ω telles que $X(\omega) \in A$.

Soit $x \in X(\Omega)$. On notera $(X = x)$ l'événement $(X \in \{x\})$ et, dans le cas où X est une variable aléatoire réelle, $(X \leq y)$ pour $(X \in]-\infty; y] \cap X(\Omega))$.

On définit de manière analogue les événements $(X \geq x)$, $(x \leq X \leq y) = (X \geq x) \cap (X \leq y) \dots$

Exemple II.3.11. (1) L'expérience aléatoire consiste à lancer 20 fois une pièce équilibrée.

On note X le nombre de "piles" obtenus. X est une variable aléatoire réelle telle que $X(\Omega)$ est fini (probabilité de première année).

Comme souvent, il faudra s'adapter au manque de rigueur des énoncés. Ici, X est bien une variable aléatoire, c'est l'**application** de Ω dans $X(\Omega) = \llbracket 0; 20 \rrbracket$ qui, à une succession de lancers, associe le nombre de Piles obtenus.

(2) L'expérience consiste à lancer deux dés à 6 faces. On note S la somme des deux chiffres obtenus.

S est une variable aléatoire réelle telle que $S(\Omega)$ est fini

(3) L'expérience consiste à lancer une pièce indéfiniment. On note X la variable aléatoire qui prend pour valeur le rang d'apparition du premier Pile et qui prend la valeur $+\infty$ si on obtient Face indéfiniment.

X est une variable aléatoire discrète et $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$. Pour tout $n \in X(\Omega)$, exprimer les événements $(X = n)$ en fonction des événements A_k et B_k déjà définis dans l'exemple (II.2.8)

III – Espaces probabilisés

1) Probabilité - Définition

Définition III.1.1.

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle probabilité sur cet espace toute application P définie sur $\mathcal{A}$:

- (1) à valeurs dans $\mathbb{R}_+$,
- (2) telle que $P(\Omega) = 1$,
- (3) telle que : pour toute famille dénombrable $(A_n)_{n \in \mathbb{N}}$ d'événements **deux à deux incompatibles** (des éléments de $\mathcal{A}$ deux à deux disjoints), la famille $(P(A_n))_{n \in \mathbb{N}}$ est sommable et on a

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n)$$

(cette propriété est appelée σ -additivité de P).

Le triplet $(\Omega, \mathcal{A}, P)$ est alors appelé espace probabilisé.

Pour tout événement $A \in \mathcal{A}$, le réel $P(A)$ est appelé probabilité de l'événement A .

Un événement est dit presque sûr s'il est de probabilité 1 ; il est dit négligeable s'il est de probabilité 0.

Proposition III.1.2. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

- (1) $P(\emptyset) = 0$;
- (2) pour toute famille **finie** $(A_i)_{1 \leq i \leq n}$ d'événements deux à deux incompatibles : $P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$;
- (3) $P(A \setminus B) = P(A) - P(A \cap B)$;
- (4) pour tous A et B dans $\mathcal{A}$: si $B \subset A$, alors $P(B) \leq P(A)$ (croissance de P) et $P(A) - P(B) = P(A \setminus B)$;
- (5) pour tout $A \in \mathcal{A}$: $P(\bar{A}) = 1 - P(A)$;
- (6) P est à valeurs dans $[0; 1]$;
- (7) pour tous A et B dans $\mathcal{A}$: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Remarque. ☺ Si l'univers Ω est **fini** (**probabilités de première année**), alors $\mathcal{A} = \mathcal{P}(\Omega)$ est fini. Dans ce cas, une suite $(A_n)_{n \in \mathbb{N}}$ d'éléments disjoints deux à deux de $\mathcal{A}$ ne contient qu'un nombre fini d'éléments non vides et on peut considérer qu'à partir d'un certain rang n_0 : $A_n = \emptyset$ et $P(A_n) = 0$. La σ -additivité s'écrit alors avec une réunion finie et une somme finie (comme le (2) du résultat précédent) et est donc équivalente à la définition d'additivité vue en première année :

$$\forall (A, B) \in \mathcal{P}(\Omega)^2, A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B).$$

Définition - Théorème III.1.3. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soit $(A_i)_{i \in I}$ une famille au plus dénombrable (finie ou dénombrable) d'événements.

- On dit que $(A_i)_{i \in I}$ est un système quasi-complet d'événements si les événements A_i sont deux à deux incompatibles et si $P\left(\bigcup_{i \in I} A_i\right) = 1$ (ou bien $\sum_{i \in I} P(A_i) = 1$).
- On dit que $(A_i)_{i \in I}$ est un système complet d'événements si les événements A_i sont deux à deux incompatibles et si $\bigcup_{i \in I} A_i = \Omega$. Cela signifie que, pour tout résultat de l'expérience, un et un seul des événements A_i est réalisé.
Par exemple $(A, \bar{A})$ est un système complet d'événements pour tout $A \in \mathcal{A}$.
- Si $(A_i)_{i \in I}$ est un système complet ou quasi-complet d'événements alors $\sum_{i \in I} P(A_i) = 1$.
La somme est une somme finie ou la somme d'une famille sommable selon la nature de I .
- Si X est une variable aléatoire discrète définie sur Ω , alors $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements.
En particulier : $\sum_{x \in X(\Omega)} P(X = x) = 1$. **EXEMPLE FONDAMENTAL!**

Exemple III.1.4. On reprend l'exemple du jeu de Pile ou Face infini.

On rappelle qu'on a convenu que $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$.

Par définition (on dit aussi *par construction*), $((X = n))_{n \in \mathbb{N}^* \cup \{+\infty\}}$ est un système complet d'événements.

2) Propriétés de deuxième année

Théorème III.5 (Continuité croissante/décroissante). Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

(1) Soit $(A_n)_{n \in \mathbb{N}}$ une suite croissante d'événements ($\forall n \in \mathbb{N}, A_n \subset A_{n+1}$). Alors la suite $(P(A_n))_{n \in \mathbb{N}}$ converge et

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n).$$

(2) Soit $(A_n)_{n \in \mathbb{N}}$ une suite décroissante d'événements ($\forall n \in \mathbb{N}, A_{n+1} \subset A_n$). Alors la suite $(P(A_n))_{n \in \mathbb{N}}$ converge et

$$P\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n).$$

Corollaire III.2.6. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $(A_n)_{n \in \mathbb{N}}$ une famille d'événements.

Les suites $\left(P\left(\bigcup_{k=0}^n A_k\right)\right)_{n \in \mathbb{N}}$ et $\left(P\left(\bigcap_{k=0}^n A_k\right)\right)_{n \in \mathbb{N}}$ convergent et :

$$(1) \quad P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=0}^n A_k\right); \quad (2) \quad P\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{k=0}^n A_k\right).$$

Exemple III.2.7. Reprenons l'exemple du jeu de Pile ou Face infini commencé en II.2.8.

Il est difficile de traiter rigoureusement une telle question sans avoir abordé l'indépendance mais nous pouvons nous appuyer sur le travail réalisé en première année en considérant que les événements A_n sont indépendants et que pour tout $n \in \mathbb{N}^*$: $P(A_n) = \frac{1}{2}$. La construction d'un espace probabilisé satisfaisant ces contraintes n'est pas un objectif du programme.

- (1) Démontrer que $P(\bar{T}) = 0$ en utilisant la suite décroissante $(\bar{B}_n)_{n \in \mathbb{N}^*}$.
- (2) Démontrer que $P(T) = 1$ en utilisant la famille des événements $((X = n))_{n \in \mathbb{N}^*}$ qui sont incompatibles deux à deux.
Nous venons de démontrer que $((X = n))_{n \in \mathbb{N}^*}$ est un système quasi-complet d'événements.
- (3) Démontrer que $P(\bar{T}) = 0$ en utilisant la famille des événements $(\bar{A}_n)_{n \in \mathbb{N}^*}$ et le corollaire précédent.

Corollaire III.2.8 (Sous-additivité). Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $(A_i)_{i \in I}$ une famille au plus dénombrable d'événements.

Si la famille $(P(A_i))_{i \in I}$ est sommable, on a :

$$P\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} P(A_i).$$

Dans le cas où $\sum_{i \in I} P(A_i) \geq 1$, cette inégalité se révèle peu intéressante ...

IV – Probabilités conditionnelles

1) Définition

Définition - Théorème IV.1.1. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit B un événement tel que $P(B) \neq 0$.

Pour tout événement A , on pose

$$P_B(A) = \frac{P(A \cap B)}{P(B)}.$$

On peut aussi noter $P(A|B)$ pour $P_B(A)$.

L'application $P_B : \mathcal{A} \rightarrow \mathbb{R}_+$ ainsi définie est une probabilité sur $(\Omega, \mathcal{A})$ appelée probabilité conditionnelle sachant B .

On dit que $P_B(A)$ est la probabilité de l'événement A sachant B .

Remarque. (1) Toutes les propriétés démontrées pour une probabilité en général sont donc valables pour P_B (à valeurs dans $[0; 1]$, probabilité de la réunion, du complémentaire, d'une réunion d'événements deux à deux incompatibles, croissance de P_B , continuité croissante ...).

(2) P est une probabilité mais, lorsqu'on écrit $P(A|B)$, on désigne la probabilité de l'événement A relativement à la probabilité conditionnelle P_B .

L'objet $(A|B)$ n'a pas d'existence propre, ce n'est surtout pas un événement!

La notation $P_B(A)$ évite toute ambiguïté.

(3)  Souvent, cette définition est utilisée sous la forme : $P(A \cap B) = P_B(A) \times P(B)$. Même si $P(B) = 0$, on pourra écrire cette égalité en adoptant la convention $P(A \cap B) = P_B(A) \times P(B) = 0$.

(4) La probabilité d'un événement $P(A)$ a été construite pour estimer la fréquence de réalisation de l'événement A si on répète un grand nombre de fois la même expérience aléatoire. La probabilité conditionnelle $P_B(A)$ est aussi construite pour estimer la fréquence de réalisation de A mais seulement parmi les expériences qui ont réalisé B . Le nombre $P_B(A)$ doit être interprétée comme la probabilité de réaliser A mais sous une condition supplémentaire : B est supposé réalisé. Il est important de comprendre cette interprétation pour pouvoir reconnaître les données de certains énoncés comme des probabilités conditionnelles.

2) Formule des probabilités composées

Théorème IV.2. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$.

Soient $(A_i)_{1 \leq i \leq n}$ une famille finie d'événements telle que

$$P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0.$$

On a alors

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \cdots \times P(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

Cette formule est appelée formule des probabilités composées.

Cette formule est particulièrement utilisée dans le cas d'une succession d'expériences aléatoires, lorsque le résultat d'une expérience dépend du résultat des précédentes.

Exercice IV.2.3. Une urne contient initialement 7 boules noires et 3 boules blanches. On tire successivement 3 boules. Lors de chaque tirage : si on tire une boule noire, on la retire de l'urne ; si on tire une boule blanche, on la retire et on la remplace par une noire. Quelle est la probabilité de tirer 3 boules blanches ?

Remarque.  L'exercice précédent est un exemple de succession d'expériences aléatoires (non indépendantes). L'univers Ω contient des listes (de 3 couleurs de boules ici). Dans ce type d'exercice, on ne cherchera (surtout) pas à expliciter Ω et la probabilité P associée. Des théorèmes techniques (et hors-programme) assurent qu'il existe une et une seule probabilité P qui vérifie les propriétés *naturellement* proposées par l'exercice :

$$P(B_1) = \frac{3}{10}, P(N_1) = \frac{7}{10}, P_{B_1}(B_2) = \frac{2}{10}, P_{N_1}(B_2) = \frac{3}{9}, \dots, P_{B_1 \cap B_2}(N_3) = \frac{9}{10}, \dots$$

Le travail de modélisation consiste alors à :

- nommer tous les événements qui seront utilisés, pour pouvoir traduire les données de l'énoncé (on pourra utiliser à profit les variables aléatoires) ;
- traduire les informations données par l'énoncé en utilisant les notations choisies à l'étape précédente : probabilités, probabilités conditionnelles, indépendances...
- traduire les questions posées en utilisant ces notations.

Reste à répondre aux questions posées en utilisant les résultats du cours !

3) Formule des probabilités totales - Formule de Bayes

Théorème IV.4. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soient I un ensemble au plus dénombrable et $(A_i)_{i \in I}$ un **système complet ou quasi-complet d'événements**.

Pour tout événement B , la famille $(P(A_i \cap B))_{i \in I}$ est finie ou sommable et on a

$$P(B) = \sum_{i \in I} P(A_i \cap B) = \sum_{i \in I} (P(A_i) \times P_{A_i}(B)).$$

Les sommes ci-dessus sont des sommes finies ou des sommes de familles sommables.

Cette formule est appelée formule des probabilités totales.

Cette formule consiste à découper l'événement B en sous-événements, selon les circonstances sous lesquelles il peut être réalisé : $B \cap A_1, B \cap A_2, \dots$

Exemple IV.3.5. On considère un lot d'objets sortis des chaînes des usines A , B et C . 8% des objets proviennent de l'usine A , 15% proviennent de l'usine B et le reste vient de l'usine C .

L'usine A produit 5% d'objets défectueux, l'usine B produit 7% d'objets défectueux et l'usine C seulement 3%.

On considère un produit sorti au hasard du lot.

- (1) Quelle est la probabilité qu'il soit défectueux?
- (2) Le produit est effectivement défectueux, quelle est la probabilité qu'il soit sorti de l'usine B ?

Exemple IV.3.6. On lance une pièce équilibrée jusqu'à obtenir Pile. Si on obtient Pile au n -ième tirage, on pioche un ticket dans une urne contenant n tickets dont un seul est gagnant.

On note X la variable aléatoire qui prend pour valeur le rang d'apparition du premier Pile et qui prend la valeur $+\infty$ si on n'obtient jamais Pile (on ne peut alors pas gagner!). On note G l'événement "piocher le ticket gagnant dans l'urne".

- (1) Calculer $P(X = n)$ pour tout $n \in \mathbb{N}^*$ et vérifier que $(X = n)_{n \in \mathbb{N}^*}$ est un système quasi-complet d'événements.
- (2) Montrer que $P(G) = \sum_{n=1}^{+\infty} \frac{1}{n2^n}$. On admet que $P(G) = \ln(2)$ (question 5/2).
- (3) Sachant que le ticket gagnant a été pioché, quelle est la probabilité que l'urne n'ait contenu qu'un seul ticket.

Proposition IV.3.7. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient A et B deux événements de probabilités non nulles. On a

$$P_B(A) = \frac{P_A(B) \times P(A)}{P(B)}.$$

Proposition IV.3.8. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit I un ensemble au plus dénombrable.

Soit $(A_i)_{i \in I}$ un système complet ou quasi-complet d'événements. Soit B un événement de probabilité non nulle.

Pour tout $k \in I$, on a

$$P_B(A_k) = \frac{P_{A_k}(B) \times P(A_k)}{\sum_{i \in I} (P(A_i) \times P_{A_i}(B))}.$$

La somme ci-dessus est une somme finie ou la somme d'une famille sommable.

On appelle formule de Bayes l'une ou l'autre de ces deux propositions.

Exercice IV.3.9. Dans une population, 1 personne sur 10000 est atteinte d'une maladie (on dit que la prévalence est de 10^{-4}). Un test de dépistage est proposé, positif à 99% chez un individu malade et positif à 0.1% chez un individu sain. Devez-vous vous inquiéter d'un patient dont le test est positif?

V – Loi d'une variable aléatoire discrète

1) Généralités - Distribution de probabilités

Définition V.1.1 (Distribution de probabilités). Soit I un ensemble au plus dénombrable.

Une famille de réels $(p_i)_{i \in I}$ est appelée distribution de probabilités si :

- pour tout $i \in I$, $p_i \geq 0$;
- la famille $(p_i)_{i \in I}$ est finie ou sommable et $\sum_{i \in I} p_i = 1$.

Exemple V.1.2. La famille finie $\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$ est une distribution de probabilités.

La famille $\left(\frac{1}{2^n}\right)_{n \in \mathbb{N}^*}$ est une distribution de probabilités.

Définition - Théorème V.1.3. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit X une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$. L'ensemble $X(\Omega)$ est alors au plus dénombrable. Munissons cet espace de la tribu $\mathcal{T} = \mathcal{P}(X(\Omega))$.

- L'application $P_X: \mathcal{T} \longrightarrow [0; 1]$ est une probabilité sur $(X(\Omega), \mathcal{T})$ appelée loi de probabilité de X .

$$A \longmapsto P_X(A) = P(X \in A)$$
- La famille $(P(X = x))_{x \in X(\Omega)}$ est une distribution de probabilités et P_X est complètement déterminée par la donnée de $X(\Omega)$ et de cette distribution de probabilités. En particulier :



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

La somme ci-dessus est une somme finie ou la somme d'une famille sommable.

Remarque. • Déterminer la loi de X revient à :

- déterminer $X(\Omega)$, c'est-à-dire l'ensemble (au plus dénombrable) des valeurs prises par X ;
- pour chaque valeur x dans $X(\Omega)$, calculer $P(X = x)$.
- La notion de *loi d'une variable aléatoire* est fondamentale puisqu'elle traduit le comportement probabiliste d'une variable aléatoire. C'est une probabilité définie sur l'univers $X(\Omega)$ des valeurs prises par la variable, univers au plus dénombrable et bien connu!

Au contraire, l'espace probabilisé $(\Omega, \mathcal{A}, P)$ est souvent difficile à expliciter et n'est pas celui que l'on rencontrera *naturellement*.

- Pour étudier une variable aléatoire, il suffit de connaître sa loi, c'est-à-dire l'ensemble $X(\Omega)$ de ses valeurs et la distribution de probabilités $(P(X = x))_{x \in X(\Omega)}$ (c'est souvent tout ce qui sera connu dans la pratique!).



Dans le cas où un énoncé présente directement une variable aléatoire par sa loi, il faudra s'assurer que la famille $(P(X = x))_{x \in X(\Omega)}$ **est bien une distribution de probabilités**.

Pour assurer l'existence d'un espace probabilisé $(\Omega, \mathcal{A}, P)$ adéquat, le théorème suivant pourrait être démontré (hors-programme) :

Soient $(\Omega, \mathcal{A})$ un espace probabilisable et X une variable aléatoire définie sur $(\Omega, \mathcal{A})$ telle que $X(\Omega)$ soit au plus dénombrable.

Soit $(p_x)_{x \in X(\Omega)}$ une distribution de probabilités. Il existe une probabilité P sur $(\Omega, \mathcal{A})$ telle que,

pour tout $x \in X(\Omega)$, $P(X = x) = p_x$.

Exemple V.1.4. Reprendre l'exemple II.3.11 et décrire les lois des différentes variables aléatoires présentées.

Définition V.1.5. Soient X et Y deux variables aléatoires discrètes.

On dit que X et Y suivent la même loi et on note $X \sim Y$ si $P_X = P_Y$

($X(\Omega) = Y(\Omega)$ et, pour tout $x \in X(\Omega)$, $P(X = x) = P(Y = x)$).

2) Lois usuelles

On rappelle les lois usuelles déjà vues en première année : loi uniforme, loi de Bernoulli et loi binomiale.

Les deux lois suivantes vont être décrites par la donnée de leur image et d'une distribution de probabilités. L'existence d'un espace probabilisé sur lequel ces variables aléatoires puissent être définies n'est pas un objectif du programme et a déjà été évoqué dans la remarque de la page précédente.

a. Loi géométrique

Définition - Théorème V.2.6. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit X une variable aléatoire discrète réelle sur $(\Omega, \mathcal{A})$.

Soit $p \in]0; 1[$.

On dit que X suit la loi géométrique de paramètre p et on note $X \sim \mathcal{G}(p)$ si :

- $X(\Omega) = \mathbb{N}^*$;
- $\forall n \in \mathbb{N}^*, P(X = n) = p(1 - p)^{n-1}$.

Remarque. ✌ Lors de la répétition (indéfiniment et de manière indépendante) d'une même épreuve de Bernoulli de paramètre p (probabilité du succès), si on note X le rang du premier succès, alors $X \sim \mathcal{G}(p)$.

Dans cette modélisation, on ne considère pas l'événement pour lequel aucune épreuve ne se solde par un succès. Ce choix est légitime puisque cet événement est négligeable. Nous l'avons déjà démontré dans le cas $p = \frac{1}{2}$.

Proposition V.2.7. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $p \in]0; 1[$.

Soit X une variable aléatoire sur $(\Omega, \mathcal{A})$ qui suit la loi géométrique de paramètre p .

Pour tout $k \in \mathbb{N}$, $P(X > k) = P(X \geq k + 1) = (1 - p)^k$.

b. Loi de Poisson

Définition - Théorème V.2.8. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit X une variable aléatoire discrète réelle sur $(\Omega, \mathcal{A})$.

Soit $\lambda \in]0; +\infty[$.

On dit que X suit la loi de Poisson de paramètre λ et on note $X \sim \mathcal{P}(\lambda)$ si :

- $X(\Omega) = \mathbb{N}$;
- $\forall n \in \mathbb{N}, P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$.

Remarque. Dans la pratique, les lois de Poisson sont utilisées pour compter des événements rares pendant un intervalle de temps fixé : nombre de mutations d'un gène, nombre de buts lors d'un match de football, nombre de bateaux arrivant dans un port ...

3) Fonction d'une variable aléatoire

Définition V.3.9. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

(1) Soit X une variable aléatoire discrète définies sur $(\Omega, \mathcal{A})$.

Soit f une fonction définie sur $X(\Omega)$. On admet que l'application $f(X): \Omega \rightarrow f(X(\Omega))$ est une variable aléatoire discrète sur $(\Omega, \mathcal{A})$.

$$\omega \mapsto f(X(\omega))$$

(2) Soit $n \in \mathbb{N}^*$. Soient $X_1, \dots, X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$.

Soit f une fonction définie sur $X_1(\Omega) \times \dots \times X_n(\Omega)$. On admet que l'application

$f(X_1, \dots, X_n): \Omega \rightarrow f(X_1(\Omega) \times \dots \times X_n(\Omega))$ est une variable aléatoire discrète sur $(\Omega, \mathcal{A})$.

$$\omega \mapsto f(X_1(\omega), \dots, X_n(\omega))$$

Exemple V.3.10.

- Si X est une variable aléatoire discrète réelle, on peut ainsi définir les variables aléatoires discrètes réelles $\sin(X)$, X^2 , λX (où $\lambda \in \mathbb{R}$),...
- Si X et Y sont deux variables aléatoires discrètes réelles, on peut définir les variables aléatoires discrètes réelles $X + Y$, XY , $\max\{X, Y\}$, $\min\{X, Y\}$,...
En particulier, l'ensemble des variables aléatoires discrètes réelles est un espace vectoriel.
- Si $X_1, X_2, \dots, X_n$ sont des variables aléatoires réelles, on peut définir les variables aléatoires discrètes réelles $\sum_{i=1}^n X_i$, $\frac{1}{n} \sum_{i=1}^n X_i$ (moyenne des X_i),...

Exercice V.3.11. 

Soient X et Y deux variables aléatoires discrètes définies sur un espace probabilisable $(\Omega, \mathcal{A})$. On suppose que $X(\Omega) = Y(\Omega) = \mathbb{N}$.

On note $M = \max\{X, Y\}$. Montrer que, pour tout $n \in \mathbb{N}$, $(M = n)$ est un événement (c'est-à-dire appartient à $\mathcal{A}$).

Ceci assure en particulier que M est une variable aléatoire.

Exercice V.3.12.

On lance un dé indéfiniment et on note X le rang du premier lancer où apparaît le premier 6. On convient que $X \sim \mathcal{G}\left(\frac{1}{6}\right)$.

On note Y la variable aléatoire qui vaut 1 si X est impair et qui vaut 0 si X est pair. Déterminer la loi de Y .

Proposition V.3.13. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$. Soit f une fonction définie sur $X(\Omega)$. Si $X \sim Y$, alors $f(X) \sim f(Y)$.

4) Couple de variables aléatoires

Définition V.4.14.

Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$. Soient deux variables aléatoires discrètes X et Y définies sur $(\Omega, \mathcal{A})$.

On appelle couple (X, Y) la variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$ par

$$\forall \omega \in \Omega, (X, Y)(\omega) = (X(\omega), Y(\omega)).$$

L'image de (X, Y) (ensemble des valeurs prises) est incluse dans le produit $X(\Omega) \times Y(\Omega)$.

Dans la pratique, quitte à considérer que certains couples sont associés à une probabilité nulle, on dit simplement que (X, Y) est à valeurs dans $X(\Omega) \times Y(\Omega)$ et on ne cherche pas à expliciter $(X, Y)(\Omega)$.

La loi conjointe de X et Y est la loi de (X, Y) .

Les lois marginales du couple (X, Y) sont les lois de X et de Y .

Remarque. • La loi de (X, Y) est complètement déterminée par la donnée de tous les couples $((x, y), P((X = x) \cap (Y = y)))$ pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$. On notera indifféremment :

$$P((X = x) \cap (Y = y)) = P((X, Y) = (x, y)) = P(X = x, Y = y).$$

-  On peut obtenir les lois marginales à partir de la loi conjointe.

En effet, $(\{Y = y\})_{y \in Y(\Omega)}$ est un système complet d'événements donc, d'après la formule des probabilités totales :

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

De même,

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

En revanche, **on ne peut pas en général déduire la loi conjointe à partir des lois marginales.**

Exemple V.4.15. On répète indéfiniment et de manière indépendante la même épreuve de Bernoulli de paramètre $p \in]0; 1[$.

On note X_1 le rang du premier succès et X_2 le rang du deuxième succès.

Pour simplifier, on considèrera que $X_1(\Omega) = X_2(\Omega) = \mathbb{N}^*$ en posant que $P(X_2 = 1) = 0$. On rappelle alors que $X_1 \sim \mathcal{G}(p)$.

(1) Déterminer la loi conjointe de (X_1, X_2) .

(2) En déduire la loi de X_2 .

Définition V.4.16. Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$. Soit X une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$.

Soit A un événement de probabilité non nulle. L'application $B \mapsto P_A(X \in B)$ de $\mathcal{T} = \mathcal{P}(X(\Omega))$ dans $[0; 1]$ définie par

$$\forall B \subset X(\Omega), P_A(X \in B) = \frac{P((X \in B) \cap A)}{P(A)}$$

est une probabilité sur $(X(\Omega), \mathcal{T})$. On l'appelle loi conditionnelle de X sachant A .

C'est la loi de la variable aléatoire X lorsqu'on munit $(\Omega, \mathcal{A})$ de la probabilité conditionnelle P_A .

Remarque. Comme toute loi de variable aléatoire discrète, la loi de X sachant A est complètement déterminée par la donnée des couples $(x, P_A(X = x))$ pour tout $x \in X(\Omega)$.

Exemple V.4.17.

On reprend les notations de l'exemple précédent. Soit $k \in \mathbb{N}^*$. Déterminer la loi conditionnelle de X_2 sachant $X_1 = k$.

Définition V.4.18.

Comme on a défini un couple de variables aléatoires discrètes, on peut définir un n -uplet de variables aléatoires discrètes.

Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$. Soit $n \in \mathbb{N}^*$.

On considère $X_1, X_2, \dots, X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$.

On définit la variable aléatoire discrète $(X_1, \dots, X_n)$ par :

$$\forall \omega \in \Omega, (X_1, \dots, X_n)(\omega) = (X_1(\omega), \dots, X_n(\omega)).$$

VI – Indépendance

1) Événements indépendants

Définition VI.1.1. Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$. Soient A et B deux événements.

On dit que A et B sont indépendants si

$$P(A \cap B) = P(A) \times P(B).$$

Lorsque la probabilité de A (respectivement de B) est non nulle, cette définition est équivalente à $P_A(B) = P(B)$ (respectivement $P_B(A) = P(A)$).

Le fait que A et B soient indépendants peut se comprendre par le fait que la probabilité que B soit réalisé n'est pas influencée par le fait que A soit réalisé ou non.

Proposition VI.1.2. Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$.

Si A et B sont deux événements indépendants, alors A et $\bar{B}$ sont indépendants.

Il en est de même pour $\bar{A}$ et B ainsi que pour $\bar{A}$ et $\bar{B}$.

2) Famille d'événements indépendants

Définition VI.2.3. Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$. Soit $n \geq 2$. Soient $A_1, A_2, \dots, A_n$ n événements.

(1) On dit que $A_1, A_2, \dots, A_n$ sont deux à deux indépendants si, pour tous i et j tels que $i \neq j$, A_i et A_j sont indépendants.

(2) On dit que $A_1, A_2, \dots, A_n$ sont indépendants si, pour tout ensemble non vide d'indices $J \subset \llbracket 1; n \rrbracket$, on a

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Remarque. (1)  Souvent, l'indépendance est inhérente à l'expérience aléatoire considérée. Elle traduit que la réalisation d'un événement de la forme $\bigcap_{j \in J} A_j$ n'influence pas la probabilité de réalisation de A_k si $k \notin J$, par exemple :

$$P_{A_1 \cap A_2}(A_3) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} = \frac{P(A_1)P(A_2)P(A_3)}{P(A_1)P(A_2)} = P(A_3).$$

C'est le cas lorsque l'expérience consiste à lancer un dé ou une pièce (ou autre chose!) n fois et que l'on note A_k un événement qui ne dépend que du k -ième tirage : les événements $A_1, \dots, A_n$ sont indépendants.

On dit simplement que **les tirages/expériences sont indépendants**.

- (2) En considérant les ensembles J à deux éléments, on remarque que si des événements sont indépendants, ils sont indépendants deux à deux.

Nous montrerons ultérieurement que la réciproque est fausse :

des événements peuvent être indépendants deux à deux et ne pas être indépendants.

Être indépendants est une propriété plus forte qu'être indépendants deux à deux.

Proposition VI.2.4. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$. Soient $A_1, \dots, A_n$ des événements indépendants. Pour tout $i \in \llbracket 1; n \rrbracket$, on note B_i l'un des événements A_i ou $\overline{A_i}$. Alors, les événements $B_1, \dots, B_n$ sont indépendants.

3) Variables aléatoires indépendantes

Définition VI.3.5. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soient X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$.

On dit que X et Y sont indépendantes lorsque pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, on a

$$P((X \in A) \cap (Y \in B)) = P(X \in A) \times P(Y \in B).$$

On note $X \perp\!\!\!\perp Y$.

- Remarque.*
- Cette définition est équivalente au fait que les événements $(X \in A)$ et $(Y \in B)$ soient indépendants pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$.
 - Intuitivement, l'indépendance de deux variables aléatoires traduit le fait que la loi de l'une n'a aucune influence sur la loi de l'autre. Rigoureusement, pour tout $A \subset X(\Omega)$ tel que $P(X \in A) > 0$, la loi conditionnelle de Y sachant $(X \in A)$ est simplement la loi de Y .

Proposition VI.3.6. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soient X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$.

Les assertions suivantes sont équivalentes :

- (1) X et Y sont indépendantes.
- (2) Pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants, c'est-à-dire

$$P((X = x) \cap (Y = y)) = P(X = x) \times P(Y = y).$$

En cas d'indépendance, la loi du couple (X, Y) (loi conjointe) se déduit donc des lois de X et Y (lois marginales).

Proposition VI.3.7. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soient X et Y deux variables aléatoires discrètes indépendantes définies sur $(\Omega, \mathcal{A})$.

Soient f et g deux applications respectivement définies sur $X(\Omega)$ et $Y(\Omega)$.

Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Exemple VI.3.8. Par exemple, si les variables aléatoires X et Y sont discrètes, réelles et indépendantes, X^2 et Y^2 sont indépendantes, X^2 et $\sin(Y)$ également,...

Exercice VI.3.9. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Soient λ et μ deux réels strictement positifs. Soient X et Y deux variables aléatoires indépendantes, que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$.

Montrer que $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.

4) Famille de variables aléatoires indépendantes

Définition VI.4.10. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$.

On considère n variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$ notées $X_1, \dots, X_n$.

On dit que les variables aléatoires $X_1, \dots, X_n$ sont indépendantes si, pour tout $(A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \dots \times \mathcal{P}(X_n(\Omega))$:

$$P\left(\bigcap_{i=1}^n (X_i \in A_i)\right) = \prod_{i=1}^n P(X_i \in A_i).$$

Remarque. Dans le cas $n = 2$, on retrouve l'indépendance de deux variables aléatoires définie dans la partie précédente.

Lemme VI.4.11. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$.

On considère n variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$ notées $X_1, \dots, X_n$.

$X_1, \dots, X_n$ sont indépendantes si et seulement si,

pour tous $x_1 \in X_1(\Omega), \dots, x_n \in X_n(\Omega)$, on a

$$P\left(\bigcap_{i=1}^n (X_i = x_i)\right) = \prod_{i=1}^n P(X_i = x_i).$$

Lemme VI.4.12. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$.

On considère n variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$ notées $X_1, \dots, X_n$.

On suppose que $X_1, \dots, X_n$ sont des variables aléatoires indépendantes.

Si on choisit une famille d'au moins 2 variables aléatoires parmi les X_i , alors ces variables aléatoires sont indépendantes.

Remarque.

En exploitant le cas des familles de 2 variables aléatoires, on déduit de ce résultat que,

si des variables aléatoires sont indépendantes, alors elle sont indépendantes deux à deux. La réciproque est fausse!

Exemple VI.4.13. On lance deux fois un dé équilibré à 6 faces. On note X_1 la valeur obtenue par le premier dé, X_2 la valeur obtenue par le second dé et X_3 qui vaut 1 si les deux dés affichent la même valeur et 0 sinon.

Les variables X_1 , X_2 et X_3 sont-elles indépendantes deux à deux? Sont-elles indépendantes?

Proposition VI.4.14 (Lemme des coalitions). Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $n \geq 2$.

Soient $X_1, \dots, X_n$ des variables aléatoires discrètes indépendantes définies sur $(\Omega, \mathcal{A})$. Soit $m \in \llbracket 1; n \rrbracket$.

Toute fonction $f(X_1, \dots, X_m)$ de $X_1, \dots, X_m$ est indépendante de toute fonction $g(X_{m+1}, \dots, X_n)$ de $X_{m+1}, \dots, X_n$.

Ce résultat se généralise au cas de plus de deux coalitions.

Par exemple : $X_1 + \dots + X_m$ et $X_{m+1} + \dots + X_n$ sont indépendantes, $X_1 X_2$, $X_3 X_4 + X_5$ et $X_6 + X_7 + \dots + X_n$ sont indépendantes ...

5) Suites de variables aléatoires indépendantes identiquement distribuées (iid)

Définition VI.5.15.

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Pour tout $n \in \mathbb{N}$, on note X_n une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$.

On dit que la suite $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes identiquement distribuées (ou suite iid) si :

- pour tout $N \in \mathbb{N}$, les variables aléatoires $X_0, X_1, \dots, X_N$ sont indépendantes;
- les variables aléatoires X_n suivent toute la même loi.

Exemple VI.5.16.  On considère la répétition (indéfiniment) d'une même épreuve de Bernoulli de paramètre $p \in [0; 1]$. C'est la situation d'un jeu de Pile ou Face infini par exemple. On modélise cette expérience en notant, pour tout $n \in \mathbb{N}^*$:

$$X_n = \begin{cases} 1 & \text{si la } n\text{-ième épreuve se solde par un succès,} \\ 0 & \text{si la } n\text{-ième épreuve se solde par un échec,} \end{cases}$$

et en supposant que $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes identiquement distribuées. En particulier, pour tout $n \in \mathbb{N}^*$, $X_n \sim \mathcal{B}(p)$.

Ce choix de modélisation a déjà été réalisé pour illustrer ce cours par l'exemple du jeu de Pile ou Face.