CHAPITRE 10 - ESPACES PROBABILISÉS

1 Univers, événéments, variables aléatoires discrètes

1.1 Généralités

Définition : Univers

Dans le contexte d'une expérience aléatoire, on appelle **univers** et on note Ω l'ensemble des issues possibles.

Remarques :

- Une issue peut être interprétée comme un état possible à la suite d'une expérience aléatoire.

  L'univers est donc un ensemble d'états que l'on peut obtenir.
- Le choix de l'univers correspond à un choix de **modélisation**.

  Différents choix de modélisations, amèneront à différents univers.

  Par exemple, si on lance deux dés cubiques à un jeu de plateau, on peut considérer l'univers des paires $[\![1,6]\!]^2$ correspondant au résultat numérique de chacun des dés après le lancer. Mais si ce qu'on veut vérifier est si l'on fait un double, il suffit d'un universe à deux éléments $\Omega = \{D, \overline{D}\}$.

  En fait, on peut aussi considérer l'univers $\Omega = \mathbb{R}^{12}$ qui donnent les 12 degrés de liberté pour décrire deux solides indéformables (position et orientation) dans l'espace.

  Le choix de l'univers doit donc se faire sur des critères de praticité et de naturalité. Il est souvent pertinent d'inclure plus de degrés de libertés que ceux stricts d'intérêts afin de pouvoir utiliser des hypothèses d'équiprobabilité par exemple. Mais il faut aussi ne pas exagérer dans l'autre sens.
- En fait, toute cette discussion va vite passer sous le tapis, car quand on aura défini les variables aléatoires, le contenu précis de Ω est souvent peu ou pas pertinent.

Définition : Tribu

Soit Ω un univers.
On appelle tribu $\mathcal{A}$ sur l'univers Ω tout ensemble de parties de Ω vérifiant :

- $\Omega \in \mathcal{A}$;
- $\forall A \in \mathcal{A},\ \overline{A} = \Omega \setminus A \in \mathcal{A}$ (stabilité par passage au complémentaire) ;
- pour toute suite d'éléments $(A_n)_n$ de $\mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ (stabilité par union dénombrable).

Le couple $(\Omega, \mathcal{A})$ est appelé espace probabilisable.

Remarques :

- Les anglo-saxons disent σ-algèbre. Certains livres français utilisent ce vocabulaire. Mais le mot au programme est bien **tribu**.
- $\mathcal{A}$ va représenter l'ensemble des événements. Un élément de $\mathcal{A}$ est donc appelé un événement.
- Intuitivement, un événement correspond à une phrase qui peut être vérifiée (ou non) sur le résultat d'une expérience aléatoire.

  Par exemple, si on lance un dé à 6 faces, on peut considérer l'univers $\Omega = [\![1,6]\!]$ et l'événement « faire un nombre pair ». Trois issues (2, 4 et 6) satisfont cette phrase. Formellement, l'événement « faire un nombre pair » sera donc l'ensemble $\{2,4,6\}$.
- L'année dernière, on ne s'était pas embêté avec la notion de tribu. En fait, un événement était, par définition, une partie de Ω.

  Cette défintion d'événement fonctionne *dans les cas simples*. Tant que Ω est fini, voir même tant que Ω est dénombrable, on peut s'en contenter.

  Malheureusement, pour des raisons techniques, dès lors que Ω est non-dénombrable, garder $\mathcal{A} = \mathcal{P}(\Omega)$ amène à des contradictions dans la définition des probabilités. On peut alors restreindre l'ensemble des événements possibles du moment que les axiomes d'une tribu sont satisfaits.

  Cela permet la cohérence de la théorie mais en contre-partie, il existe potentiellement des parties de Ω qui ne sont pas des événements.
- Cela signifie en particulier que Ω est un événement. On l'appelle l'événement certain.

  De même si A est un événement, l'événement contraire $\overline{A}$ est aussi un événement.

  L'union dénombrable nous assure que l'on peut construire des événements à partir de « ou » puisque $A \cup B$ signifie « A ou B ». Et en combinant avec les événements contraires, on peut aussi faire apparaître la conjonction « et » puisque $\overline{A} \cap \overline{B} = \overline{A \cup B}$.
- On peut montrer que $\varnothing \in \mathcal{A}$ puisque $\varnothing = \overline{\Omega}$. On appelle cette événement, l'événement impossible.

Exemples :

- La tribu triviale (ou grossière) : $\{\varnothing, \Omega\}$.

- La tribu pleine (ou discrète) : $\mathcal{P}(\Omega)$ (qui ne fonctionne pour la suite que si Ω est au plus dénombrable).
- La tribu engendré par un événement : $\{\varnothing, A, \overline{A}, \Omega\}$.

1.2 Opérations sur les événements

Définition

On dit qu'un événement $A \in \mathcal{A}$ est réalisé si l'issue $\omega \in \Omega$ réalisée est dans A.

Remarques :

- Cela explique l'appelation *événement contraire* puisque :

$$\left(\omega \in \overline{A}\right) \Leftrightarrow \left(\omega \notin A\right).$$

- Cela explique les liens entre les opérations d'union et la conjonction « ou ». En effet :

$$\left(\omega \in A \cup B\right) \Leftrightarrow \left(\omega \in A \text{ ou } \omega \in B\right).$$

- De même :

$$\left(\omega \in A \cap B\right) \Leftrightarrow \left(\omega \in A \text{ et } \omega \in B\right).$$

- De même on a :

$$\left(\omega \in \bigcup_{n=0}^{+\infty} A_n\right) \Leftrightarrow \left(\exists n \in \mathbb{N},\ \omega \in A_n\right).$$

Donc $\bigcup_{n=0}^{+\infty} A_n$ est réalisé si et seulement si **l'un des** A_n est réalisé.

Proposition : Stabilité par intersection dénombrable

Soit $\mathcal{A}$ une tribu sur Ω et soit (A_n) une suite d'événéments de $\mathcal{A}$. On a :

$$\bigcap_{n=0}^{+\infty} A_n \in \mathcal{A}.$$

Démonstration : ***À faire en classe.*** $\square$

Remarque : encore une fois, on a :

$$\left(\omega \in \bigcap_{n=0}^{+\infty} A_n\right) \Leftrightarrow \left(\forall n \in \mathbb{N},\ \omega \in A_n\right).$$

Donc $\bigcap_{n=0}^{+\infty} A_n$ est réalisé si et seulement si **tous les** A_n sont réalisés.

Définition

On dit que A et B sont incompatibles si $A \cap B = \varnothing$.

Remarque : $\varnothing$ est l'événement impossible. On dit que « A et B » est un événément impossible, c'est-à-dire qu'il n'est pas possible d'avoir A et B simultanément.

1.3 Variables aléatoires

Définition : Variable aléatoire discrète

Soit $X : \Omega \to X(\Omega)$ une application sur Ω.
On dit que X est une variable aléatoire discrète si $X(\Omega)$ est au plus dénombrable et si pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement.

Exemple : Si A est un événement, alors $\mathbb{1}_A$ est une variable aléatoire discrète.
Remarques :

- Commençons par décrypter pourquoi on considère une fonction sur Ω.
  X est une fonction qui, à chaque issue, associe un résultat. Cela encode le fait que la valeur de X dépend de l'issue obtenue.
  Ainsi, tant qu'on travaille dans le monde abstrait des probabilités, X est bien une fonction qui associe à chaque potentiel résultat d'une expérience aléatoire, une valeur observable.
  Mais, si on faisait effectivement l'expérience, on ne trouverait qu'une seule valeur pour X, celle effectivement obtenue pour l'issue ω réalisée $(X(\omega))$.
- Une variable aléatoire est donc en fait une fonction déterministe. On dit parfois que les variables aléatoires sont comme les cochons d'Inde, qui ne sont pas des cochons, et qui ne viennent pas d'Inde.
- En général, on considérera des variables aléatoires *réelles* (éventuellement complexes) donc avec $X(\Omega) \subset \mathbb{K}$.
  Mais, *a priori*, on peut imaginer associer des tas de choses aux issues. Par exemple, si on tire dans une urne des boules colorées et numérotés, une variable aléatoire pourrait être le numéro de la boule tirée. Ce serait donc une variable aléatoire réelle. Mais on pourrait aussi considérer la couleur tirée, et la variable ne serait alors pas réelle.
- On a défini les variables aléatoires *discrètes*. Cette définition contient *a priori* deux morceaux : le morceau « variable aléatoire » et le morceau « discrète ».
  Le morceau facile à définir est « discrète ». Si X est une variable aléatoire, X est discrète si et seulement $X(\Omega)$ est (au plus) dénombrable. En revanche, le morceau « variable aléatoire » est étonnamment difficile à définir seul, c'est d'ailleurs hors-programme.

Sans rentrer dans le détail, essayons d'expliquer l'origine de la difficulté. Soit $X : \Omega \to X(\Omega)$ une variable aléatoire (on imagine que ça a un sens). On peut alors considérer l'expérience aléatoire qui consiste à regarder les résultats de X. Dit autrement, on peut considérer une *expérience aléatoire dérivée* dont les valeurs de X constitue l'universe probabiliste.

Mais, si X n'est pas discrète, cela signifie que $X(\Omega)$ n'est pas au plus dénombrable. On a donc $\Omega' = X(\Omega)$ non dénombrable. Quelle tribu peut-on alors poser sur Ω' ? On a vu que $\mathcal{P}(\Omega')$ ne fonctionne pas dans ce cas. C'est là, la difficulté.

Ici, on contourne la difficulté en ne regardant que des variables aléatoires discrètes. En conséquences, $\Omega' = X(\Omega)$ est au plus dénombrable et donc $\mathcal{P}(\Omega')$ est une tribu tout à fait raisonnable.

Proposition

Soit $V \subset X(\Omega)$. $X^{-1}(V) = \{\omega \in \Omega,\ X(\omega) \in V\}$ est un événement.

Démonstration : ***À faire en cours.*** □

Remarques :

- Cela signifie que l'on peut définir un événement à partir des valeurs prises par une variable aléatoire X. Puisque cela arrive assez souvent, on pose des notations associées.

  On notera $(X = x)$ l'événement $X^{-1}(\{x\})$. On trouve aussi parfois la notation $\{X = x\}$ (et aussi $[X = x]$ qui est celle que j'ai apprise étant plus jeune mais n'est plus au programme...).

  De même, on notera $(X \in V) = X^{-1}(V)$.
- Ces notations correspondent à l'intuition derrière les événements. Par exemple, pour un lancer de dé, si X donne la valeur numérique du dé, alors $(X = 6)$ est l'événement « faire un 6 » et $(X \in 2\mathbb{N})$ est l'événement « faire un nombre pair ».
- Si X est à valeurs réelles, alors $(X \in]-\infty, x])$ est un événement. On peut le noter plus simplement $(X \leqslant x)$. De même, on peut définir $(X \geqslant x)$, $(X < x)$ et $(X > x)$.

Proposition

Si X et Y sont des variables aléatoires à valeurs dans un même ensemble, alors $\{\omega \in \Omega,\ X(\omega) = Y(\omega)\}$ est un événement.

Démonstration : ***À faire en cours.*** □

Remarque : Encore une fois, on pose une notation associée plus simple à manipuler $(X = Y)$.

2 Probabilités

2.1 Généralités

Définition : Probabilité sur un espace probabilisable

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega, \mathcal{A})$ toute fonction $\mathbb{P}$ définie de $\mathcal{A}$ dans $[0, 1]$ vérifiant :

- $\mathbb{P}(\Omega) = 1$;
- pour toute suite $(A_n)_n$ d'événements deux à deux incompatibles :

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).\ (\sigma\text{-additivité})$$

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

Remarques :

- On trouve la notation P (au programme) et la notation $\mathbb{P}$ (dans les livres et les sujets).

  La probabilité d'un événement A sera donc noté $P(A)$ ou $\mathbb{P}(A)$.
- Si $(A_k)_{k \in [\![0,n]\!]}$ est une suite d'événements deux à deux incompatibles, on peut encore appliquer la σ-additivité. En effet, si on pose $A_k = \varnothing$ si $k > n$, la suite $(A_k)_{k \in \mathbb{N}}$ est alors une suite infinie d'événements deux à deux incompatibales. On a donc :

$$\mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right) = \sum_{k=0}^{n} \mathbb{P}(A_k).$$

- En particulier, si $\Omega = \{x_1, \dots, x_n\}$ est fini, alors on peut calculer toutes les probabilités à partir des $\mathbb{P}(\{x_i\})$. En effet, on a :

$$P(A) = \sum_{x \in A} \mathbb{P}(\{x\}).$$

- Encore une fois, cela permet d'éclairer la définition et sa complexité. *A priori*, l'intuition voudrait qu'on définisse la probabilité des issues de Ω puis qu'on généralise aux événements. D'ailleurs, ça marche dans le cas d'un univers fini.

  Mais, que dire si Ω n'est pas dénombrable ? Dans ce cas, on ne peut plus facilement assembler les événements à partir d'un nombre au plus dénombrable d'issues et le calcul ne fonctionne pas.

  C'est pourquoi la définition est si compliqué : elle est valable même dans les cas délicat où Ω n'est pas dénombrable.

Définition

Soit A un événement. On dit que A est presque sûr (ou presque certain) si $P(A) = 1$. De même, on dit que A est négligeable (ou quasi-impossible) si $P(A) = 0$.

Remarques :

- Soit (A_n) une famille d'événements deux à deux disjoints et non impossibles. On dit que (A_n) est un système complet d'événements si $\bigcup_{n=0}^{+\infty} A_n = \Omega$.
  On peut modifier légèrement la définition. On dit que (A_n) est un système quasi-complet d'événements si $\sum_{n=0}^{+\infty} P(A_n) = 1$. Par σ-additivité, cela revient à dire que $\bigcup_{n=0}^{+\infty} A_n$ est presque sûr.
- Soit (A_n) un système quasi-complet d'événements mais pas un système complet. Posons $B = \bigcup_{n=0}^{+\infty} A_n$. On a donc $B \neq \Omega$. Alors $(\overline{B}, A_0, A_1, \ldots)$ est une système complet d'événements.

2.2 Propriétés

Proposition

On a :

- $P(\overline{A}) = 1 - P(A)$;
- $P(\varnothing) = 0$;
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- $P(A \cup B) \leqslant P(A) + P(B)$;
- $P(A) \leqslant P(B)$ si $A \subset B$ (croissance) ;
- $P(A \setminus B) = P(A) - P(B)$ si $B \subset A$.

Démonstration : ***À faire en classe.*** □

Remarque : A est presque sûr si et seulement si $\overline{A}$ est négligeable.

Proposition : Limite monotone

Pour toute suite (A_n) d'événements croissante, on a :

$$\lim_{n \to +\infty} P(A_n) = P\left(\bigcup_{n=0}^{+\infty} A_n\right).$$

De même pour toute suite (A_n) d'événements décroissante, on a :

$$\lim_{n \to +\infty} P(A_n) = P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

Démonstration : ***À faire en classe.*** □

Corollaire

Soit (A_n) une suite d'événements (pas nécessairement monotone). On a :

$$\lim_{n \to +\infty} P\left(\bigcup_{k=0}^{n} A_k\right) = P\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

De même, on a :

$$\lim_{n \to +\infty} P\left(\bigcap_{k=0}^{n} A_k\right) = P\left(\bigcap_{k=0}^{+\infty} A_k\right).$$

Proposition : Sous-additivité

Pour toute suite (A_n) d'événements (pas nécessairement incompatibles entre eux), on a :

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{k=0}^{+\infty} P\left(A_k\right).$$

Remarque : on pose $\sum_{k=0}^{+\infty} P\left(A_k\right) = +\infty$ si la série diverge, ainsi la proposition reste vraie même dans ce cas.

3 Probabilités conditionnelles

3.1 Généralités

Définition

Soit B un événement avec $P(B) > 0$.
On appelle probabilité de A sachant B la quantité :

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

Remarques :

- Il faut évidemment que $P(B) > 0$ pour que la fraction ait un sens.
  Mais on peut comprendre cette condition autrement. Le but de cette définition est de correspondre à l'intuition suivante : sachant que B est réalisé, quelle est désormais la probabilité que A le soit aussi ? Et quel sens pourrait-on donner à « B est réalisé » si $P(B) = 0$?
- On a donc $P(A \cap B) = P_B(A)P(B)$. Et ensuite :

$$P_B(A)P(B) = P(A \cap B) = P_A(B)P(A).$$

Proposition

L'application $A \mapsto P_B(A)$ définit une probabilité sur $(\Omega, \mathcal{A})$.

Démonstration : ***À faire en cours.*** □

3.2 Divers formules

Proposition : Formules des probabilités composées

Soit $(A_1, \ldots, A_n)$ une famille d'événements. Si $P\left(\bigcap_{k=0}^{n-1} A_k\right) \neq 0$ alors :

$$P\left(\bigcap_{k=0}^{n} A_k\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap \cdots A_{n-1}}(A_n).$$

Démonstration : ***À faire en cours.*** □

Remarque : Si $A_n \subset A_{n-1} \subset A_{n-2} \subset \cdots \subset A_1$ alors :

$$P\left(\bigcap_{k=0}^{n} A_k\right) = P(A_1)P_{A_1}(A_2)P_{A_2}(A_3) \cdots P_{A_{n-1}}(A_n).$$

Proposition : Formule des probabilités totales

Soit (A_n) un système complet ou quasi-complet dénombrable d'événements. On a :

$$P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n)$$

où on pose $P_{A_n}(B)P(A_n) = 0$ si $P(A_n) = 0$.

Démonstration : ***À faire en cours.*** □

Exemple : il y a deux urnes. La première urne contient 3 boules noires et 2 boules blanches. La seconde contient 5 boules noires et 1 boule blanche. On choisit une urne au hasard, puis on pioche une boule. Quelle est la probabilité d'obtenir une boule noire ?

Proposition : Formule de Bayes

Si A et B sont deux événements de probabilités non nulles, alors :

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

Si (A_n) est un système complet ou quasi-complet d'événements de probabilités non nulles alors :

$$P_B(A_j) = \frac{P(A_j)P_{A_j}(B)}{P(B)} = \frac{P(A_j)P_{A_j}(B)}{\sum_{n=0}^{+\infty} P(A_n)P_{A_n}(B)}$$

Exemple : un test permet de détecter la séropositivité à un virus avec une fiabilité de 99% (aussi bien pour le risque de faux positif que de faux négatif). La maladie touche 10% de la population. Si je suis testé positif, quelle est la probabilité que je suis effectivement malade ?

4 Indépendance

4.1 Indépendances de deux événements

Définition : Indépendances de 2 événements

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient $A, B \in \mathcal{A}$ 2 événements. On dit que A et B sont indépendants si :

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

Proposition

Si $P(A) \neq 0$, A et B sont indépendants si et seulement si $P_A(B) = P(B)$.

Proposition

Si A et B sont indépendants alors A et $\overline{B}$ sont indépendants.

Démonstration : ***À faire en cours.*** □

4.2 Indépendances de n événements

Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient $A_1, \ldots, A_n \in \mathcal{A}$ n événements. On dit que :

- $A_1, \ldots, A_n$ sont deux à deux indépendants si pour tout $(i, j) \in [\![1, n]\!]^2$ avec $i \neq j$, A_i et A_j sont indépendants c'est-à-dire :

$$\forall (i, j) \in [\![1, n]\!]^2,\ i \neq j \Rightarrow P(A_i \cap A_j) = P(A_i)P(A_j).$$

- $A_1, \ldots, A_n$ sont mutuellement indépendants si on a :

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

Remarques :

- $(i, j) \in [\![1, n]\!]^2$ avec $i \neq j$ définit une sous-partie à 2 éléments de $[\![1, n]\!]$. Donc si n événements sont mutuellement indépendants, ils sont *a fortiori* deux à deux indépendants.
- En revanche, des événements peuvent être deux à deux indépendants sans être mutuellement indépendants.

  Par exemple : on lance deux pièces de monnaies. On définit :

  - A : la pièce 1 tombe sur pile ;
  - B : la pièce 2 tombe sur pile ;
  - C : les deux pièces tombent sur la même face.

  A, B, C sont deux à deux indépendants mais pas mutuellement indépendants.
- Une façon de comprendre l'indépendance est de le voir comme une question : « est-ce que connaître le résultat de certains événements m'apportent des informations sur un autre ? » Si la réponse est toujours **non** alors c'est mutuellement indépendant.

  Si la réponse est **non** dans le cas réduit « est-ce qu'un événement me donne des informations sur un des autres » alors c'est deux à deux indépendants.

  Le voir avec l'exemple précédent.

Proposition

Si A_1, ..., A_n sont mutuellement indépendants alors A_1, ..., A_{n-1} et $\overline{A_n}$ sont indépendants.

Remarque : on peut arbitrairement passer certains des événements à l'événement contraire.