Chapitre 3

Probabilités (Généralités)

Version allegée

Pour éviter tout excès de technicité, certaines définitions utilisées dans ce chapitre sont un peu édulcorées par rapport à celles véritablement utilisées par les probabilistes. Elles sont cependant à notre programme telles que données ici.

I Le triplet fondamental

I-1 Cas des probabilités finies

I.1-a) Rappel de définition et propriétés élémentaires

Notation : Pour un ensemble Ω, on note $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω :
$$\mathcal{P}(\Omega) = \{A \mid A \subset \Omega\}$$

Remarque :

Si Ω est de cardinal fini n, on a card$(\mathcal{P}(\Omega))$ $=2^n$.

Définition (rappel)

Pour tout ensemble **fini** non vide Ω, on appelle *mesure (ou loi) de probabilité* p sur $(\Omega, \mathcal{P}(\Omega))$ toute application $p : \mathcal{P}(\Omega) \to \mathbb{R}$ telle que

- i ■ $\forall A \in \mathcal{P}(\Omega)$, on a $0 \leqslant p(A) \leqslant 1$
- ii ■ $p(\emptyset) = 0\,; \quad p(\Omega) = 1$
- iii ■ Pour tout couple (A, B) d'éléments incompatibles de $\mathcal{P}(\Omega)$, on a

$$p(A \sqcup B) = \text{p(A)} + \text{p(B)}$$

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

Étant donné un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), p)$, on rappelle également le vocabulaire suivant :

- on appelle Ω l'*ensemble fondamental* (ou l'*univers*).
- Un élément de Ω est appelé *épreuve* (ou événement fondamental.)
- Un élément de $\mathcal{P}(\Omega)$ est appelé *évènement.*

Exercice 1

Montrer que la condition $p(\emptyset) = 0$ de la définition est redondante.

Solution

On a $\Omega = \emptyset \cup \Omega$ une réunion d'ensembles incompatibles. Ainsi, on a
$$p(\Omega) = p(\emptyset) + p\Omega)$$
Il suffit d'isoler $p(\emptyset)$ d'un côté de l'égalité pour obtenir directement
$$p(\emptyset) = p(\Omega) - p(\Omega) = 0$$

Propriété

On rappelle également que la propriété iii de la définition nous amène à un résultat plus général qui est le suivant :
Si $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$ sont des ensembles deux à deux incompatibles, alors
$$p(A_1 \sqcup \ldots \sqcup A_n) = p(A_1) + \ldots + p(A_n)$$

■ **Exemple 1** :

Une compagnie pétrolière dispose de 4 bateaux numérotés de 1 à 4. Les bateaux 3 et 4 sont plus récents. L'expérience est la suivante : pour faire un voyage, le compagnie choisit un des bateaux, avec une probabilité deux fois plus importante de choisir un bateau récent qu'un bateau ancien.
On modélise cette expérience par le triplet fondamental $(\Omega, \mathcal{P}(\Omega), p)$ suivant :
$\Omega = \{1, 2, 3, 4\}$ p décrit par

i	1	2	3	4
$p(\{i\})$	a	a	2a	2a

où $a \in [0, 1]$. Pour trouver a, il nous suffit d'utiliser la propriété ci-dessus avec les 4 ensembles 2 à 2 incompatibles $A_i = \{i\}$ pour $i \in \Omega$. Ainsi :
$$a + a + 2a + 2a = p(A_1) + \ldots + p(A_4) = p(A_1 \cup \ldots \cup A_4) = p(\Omega) = 1$$
D'où
$$a = \frac{1}{6}$$
On peut également trouver de la même manière par exemple
$$p(\{1, 3\}) = p(\{1\}) + p(\{3\}) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}.$$

Remarque :

Généralement, à des fins de simplicité d'écriture, on préfère utiliser <u>l'abus de notation</u> $p(i)$ à la place de $p(\{i\})$ si $i \in \Omega$.

I.1-b) Problématiques, limites et insuffisances de la théorie de l'univers fini

Première problématique : connait-on Ω ?

Pour une expérience donnée, on a deux possibilités pour l'univers :
- **on connait à l'avance** exactement l'ensemble des résultats possibles *(par exemple, un lancé de dé à 6 faces)*
- **on ne connait pas à l'avance** l'ensemble des résultats *(par exemple, si on étudie l'ensemble les effets secondaires d'un médicament)*

Dans le premier cas, on peut énoncer clairement un univers modélisant l'expérience *(par exemple, $\Omega = \{1, \ldots, 6\}$)*.
Dans le deuxième cas, on sait simplement qu'il existe. On ne connait pas ses éléments et on ne sait pas du tout combien il en contient. On doit donc se contenter d'exploiter ce qu'on sait.

Deuxième problématique : Peut-on se contenter d'Ω fini ?

Faisons une expérience simple :

On lance une pièce de monnaie et on s'arrête dès que l'on obtient "pile".

On s'intéresse dans un premier temps aux résultats exacts de l'expérience.
On peut par exemple représenter chaque épreuve par une série de lettres P et F qui représentent respectivement l'obtention d'un pile ou d'un face. (FFFP représenterait donc l'obtention de trois "faces" avant l'obtention du premier pile.) On a

$$\Omega = \{P, FP, F \ldots FP \mid \text{la longueur de } F \ldots F \text{ est quelconque.}\}.$$

L'univers ainsi constitué peut-il être fini ? Non. *(à méditer)*. **Peut-on construire un univers fini pour décrire l'ensemble des résultats de cette expérience ?** Non. *(à méditer)*.

On s'intéresse dans un second temps plus simplement au rang d'apprition du premier pile. Par exemple, ce nombre serait "4" dans le cas de "FFFP"). En modélisant par 0 si Pile n'apparait jamais, L'univers obtenu est

$$\Omega = \mathbb{N}$$

infini là aussi...

Troisième problématique : peut-on appliquer la définition d'une mesure de probabilité finie a à Ω infini ?

La réponse à cette question est là aussi : Non. Pour plusieurs raisons :

1. Avec la définition "basique" donnée jusque là, on ne peut pas calculer la probabilité d'une réunion infinie d'événements.
2. Se contenter d'utiliser comme ensemble d'événements $\mathcal{P}(\Omega)$ se révèle en réalité trop complexe pour l'étude des probabilités en général, par exemple dans le cas où $\Omega = \mathbb{R}$ qui contient trop de sous-ensembles "compliqués" (et HP) qu'il est inutile de prendre en compte lors de calculs de probabilités.

   Cet aspect technique de construction d'univers sera très brièvement évoqué cette année. (*cf* ci-après.)

I-2 Cas général : mesure de probabilité

I.2-a) Réunions et intersections "dénombrables"

Commentaires :

Un univers constitué d'une infinité d'éléments doit certainement pouvoir se "découper" en une infinité d'ensembles !
Par exemple, $\Omega = \mathbb{N}$ est constitué d'une infinité de petits morceaux qui sont les $\{i\}$ pour tout entier naturel i. Par le principe de réunion, vient alors naturellement l'idée d'écrire

$$\mathbb{N} = \bigcup_{i=0}^{+\infty} \{i\}$$

C'est une réunion "infinie" qu'il nous faut donc définir mathématiquement (on en profite pour décrire également les intersections !) :

Définition

Les opérateurs de réunion et d'intersection infinie ont le même sens que dans le cadre fini :

1. L'événement $\bigcup_{n=1}^{+\infty} A_n$ correspond à *"au moins un des événements A_n se produit."*
2. L'événement $\bigcap_{n=1}^{+\infty} A_n$ correspond à *tous les événements A_n se produisent.*

Une réunion (resp. intersection) avec des indices dans $\mathbb{N}$ s'appelle une *réunion (resp. intersection) dénombrable*.

Remarque :

Pour ce qui est des probabilités, si on a un univers $\Omega = \mathbb{N} = \bigcup_{i=0}^{+\infty} \{i\}$, on aura alors très rapidement envie *(oui, oui !)* d'écrire par exemple

$$p(\mathbb{N}) = \sum_{i=0}^{+\infty} p(i)$$

mais pour l'instant, la définition de probabilités finies ne nous le permet pas. (On peut faire seulement des sommes finies.)

I.2-b) Notion de tribu

■ Exemple 2 :

Imaginons par exemple l'expérience suivante :
On lance deux fléchettes sur une cible carré lontaine de côté 1 mètre. La distance entre les deux impacts nous amène alors à considérer l'univers $\Omega = [0, \sqrt{2}]$ *(unité en mètre)*. On a donc un univers qui n'est ni fini, ni $\mathbb{N}$, mais un intervalle de $\mathbb{R}$.

Commentaires :

Dans le cadre de probabilités en général, on souhaite pouvoir :
- calculer des probabilités de réunions (finies ou dénombrables)
- calculer des probabilités d'intersections (finies ou dénombrables)
- calculer des probabilités de complémentaires
- dire que $p(\Omega) = 1$
- et si $\Omega = \mathbb{R}$ *(ou tout intervalle de* $\mathbb{R}$*), on veut pouvoir calculer des probabilités d'intervalles (ouverts, fermés, semi-ouverts)*

Si on conserve la définition

$$p : \mathcal{P}(\Omega) \to [0; 1]$$

(Non, le domaine de définition d'une probabilité n'est pas Ω *!)*
La généralisation du reste de la définition devrait prendre en compte le fait que $\mathbb{R}$ *contienne des sous-ensembles "compliqués", qui ne peuvent par exemple pas s'écrire sous une forme de réunion ou d'intersection dénombrable. (exemple HP)*
Plutôt que de compliquer inutilement cette définition en tenant compte d'ensembles auxquels ont ne s'intéresse pas en calcul de probabilités, on va plutôt restreindre le domaine de définition de p *à des cas "plus simples à étudier" :*

Imaginons donné un triplet $(\Omega, \mathcal{T}, p)$*, avec* $\mathcal{T} \subset \mathcal{P}(\Omega)$*. Si* $p : \mathcal{T} \to [0; 1]$ *doit être une probabilité, il est clair qu'elle doit au moins conserver les propriétés de base d'une probabilité finie ainsi que celles évoquées ci-dessus. On doit donc au moins avoir*

- $\emptyset \in \mathcal{T}$ *et* $\Omega \in \mathcal{T}$ *;*
- $A \in \mathcal{T} \quad \Rightarrow \quad \bar{A} \in \mathcal{T}$ *;*
- $A, B \in \mathcal{T} \quad \Rightarrow \quad A \cup B \in \mathcal{T}$*. Plus généralement* $A_1, \ldots, A_n, \ldots \in \mathcal{T} \quad \Rightarrow \quad \bigcup_{n=1}^{+\infty} A_n \in \mathcal{T}$.
- $A, B \in \mathcal{T} \quad \Rightarrow \quad A \cap B \in \mathcal{T}$*. Plus généralement* $A_1, \ldots, A_n, \ldots \in \mathcal{T} \quad \Rightarrow \quad \bigcap_{n=1}^{+\infty} A_n \in \mathcal{T}$.

En retirant les conditions redondantes, on obtient la définition suivante :

Définition

Soit Ω un ensemble fondamental et $\mathcal{T} \subset \mathcal{P}(\Omega)$. On dit que $\mathcal{T}$ est une *tribu* sur Ω si

- $\Omega \in \mathcal{T}$,
- $\mathcal{T}$ est stable par passage au complémentaire,
  (i.e. si $A \in \mathcal{T}$, alors $\bar{A} \in \mathcal{T}$.)
- $\mathcal{T}$ est stable par réunion dénombrable :
  (i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{T}$, on a $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{T}$.)

Propriété

Si $\mathcal{T}$ est une tribu, alors elle est également stable par intersection.

`Démonstration :` admise. □

■ Contre-Exemple(s) :

3 ■ Si $\Omega = \{1, 2, 3, 4\}$, l'ensemble $\{\emptyset, \Omega, \{1\}, \{1, 2\}\}$ n'est pas une tribu.
4 ■ Si $\Omega = \mathbb{Z}$, l'ensemble $\{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \ldots\}$ n'est pas une tribu.

■ Exemples :

Soit Ω un ensemble fondamental.
5 ■ $\{\emptyset, \Omega\}$ est une tribu.
6 ■ $\mathcal{P}(\Omega)$ est une tribu quelquesoit Ω.
7 ■ Si $\Omega = \{1, 2, 3, 4\}$, une tribu sur Ω peut être $\mathcal{T} = \{\emptyset, \Omega, \{1, 2\}, \{3, 4\}\}$
8 ■ Si $\Omega = \mathbb{N}$, une tribu sur Ω peut être $\mathcal{T} = \{\emptyset, \Omega, 2\mathbb{N}, 2\mathbb{N} + 1\}$

Définition

Pour tout ensemble Ω et toute tribu $\mathcal{T}$ sur Ω, on appelle le couple $(\Omega, \mathcal{T})$ un *espace probabilisable* (ou *espace mesurable*).

I.2-c) La tribu des Boréliens

Commentaires :

Si l'ensemble fondamental est $\mathbb{R}$ *ou un intervalle de* $\mathbb{R}$*, au vue des raisons évoquées précédemment, on choisit donc généralement une tribu plus petite que* $\mathcal{P}(\Omega)$ *:*

Définition et Proposition

Si l'univers $\Omega \subset \mathbb{R}$ est un intervalle, il existe une tribu ayant les propriétés suivantes :
- elle contient l'ensemble $\mathcal{C} = \{[a, b] \mid a, b, \in \Omega\}$
- il n'existe pas de tribu plus petite contenant $\mathcal{C}$
On la note $\mathcal{B}(\Omega)$. On l'appelle *tribu des boréliens* sur Ω.

Concrètement :
On peut montrer **(et on va simplement retenir)** que cette tribu fait tout ce dont on a besoin dans $\mathbb{R}$: elle contient les intervalles de tout type (fermés, bornés, semi-ouverts) leurs réunion, intersections, complémentaires, etc...
De plus, conformément à ce qui a évoqué en début de paragraphe, $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$! *(démo toujours HP)*

I.2-d) Espace probabilisé

Commentaires :
Tout d'abord, on récupère le vocabulaire des espaces finis :

Définition
Étant donné un espace probabilisé $(\Omega, \mathcal{T}, p)$,

- on appelle Ω l'*ensemble fondamental* (ou l'*univers*).
- Un élément de Ω est appelé *épreuve.*
- Un élément de $\mathcal{T}$ est appelé *évènement.*
- Pour un évènement $A \in \mathcal{T}$, l'événement $\bar{A} = A^c$ est appelé *évènement contraire de A.*
- Si deux événements A et B sont disjoints, alors ils sont dits *incompatibles*.

Commentaires :
Ensuite on généralise la notion de mesure de probabilités :

Définition
Étant donné un ensemble Ω et une tribu $\mathcal{T}$ sur Ω, on appelle *mesure de probabilité* p sur $(\Omega, \mathcal{T})$ toute application $p : \mathcal{T} \to \mathbb{R}$ telle que

1. $\forall A \in \mathcal{T}$, on a $0 \leqslant p(A) \leqslant 1$
2. $p(\emptyset) = 0\,; \quad p(\Omega) = 1$
3. *axiome de $\Sigma-$additivité :*

   Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles de $\mathcal{T}$, la série $\sum_{n\in\mathbb{N}} p(A_n)$ est convergente et $p(\bigsqcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} p(A_n)$.

Le triplet $(\Omega, \mathcal{T}, p)$ est alors appelé *espace probabilisé*.

Remarque :

- Si on a $A_n = \emptyset$ (et donc $p(A_n) = 0$) pour tout $n > N$, on obtient alors, **sous réserve d'incompatibilité des** A_n : $p\left(\bigsqcup_{n=0}^{N} A_n\right) = \sum_{n=0}^{N} p(A_n)$.
- Si on suppose Ω fini, $\mathcal{T} = \mathcal{P}(\Omega)$ et que l'on se donne $A_n = \emptyset\ \forall n \geqslant 2$ (et donc $p(A_n) = 0$) dans la troisième hypothèse, on **retrouve la définition d'une mesure de probabilité finie** donnée en première année.

Remarque :

- **L'ordre dans lequel on fait la somme des $p(A_n)$ n'importe pas.** On rappelle en effet que nous avons vu dans le chapitre sur les séries que si la série est convergente et de termes positifs (ce qui est le cas ici), on pouvait sommer dans n'importe quel ordre.
- **Si les A_n sont bien incompatibles**, la convergence de la série $\sum_{n\in\mathbb{N}} p(A_n)$ est en réalité une **hypothèse redondante**. En effet, on sait qu'elle est forcément convergente car c'est une série à termes positifs et majorée (par 1) :
$$\forall n \in \mathbb{N}, \qquad \sum_{k=0}^{n} p(A_k) = p\left(\bigsqcup_{k=0}^{n} A_k\right) \leqslant 1$$
- Cette année, les tribus utilisées seront, soit $\mathcal{P}(\Omega)$ (probabilités finies ou discrètes), soit la tribu des boréliens (probabilités à densité.)

Commentaires :
Observons maintenant l'émergence de nouveaux type d'événements :

Exemple 9 :
Choisissons au hasard un nombre dans l'intervalle $[0, 1]$. Intuitivement, réfléchir à la probabilité possible de l'événement A : "Obtenir 0" ainsi qu'à l'événement : B : "Obtenir un nombre contenu dans l'intervalle $]0, 1[$".

On ne démontrera pas encore les choses ici mais on peut imaginer que si on généralise **(de manière totalement scandaleuse !)** le principe d'équiprobabilité vu dans le cadre des probas finies, on devrait avoir
$$p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre total de cas}} = \text{« } \frac{1}{\infty} \text{ »} = 0$$
Même si on n'est pas encore en capacité de démontrer les choses proprement, on voit mal quel pourrait être une autre probabilité pour cet événement. **On a donc un événement de probabilité nulle, mais qui n'est pas impossible !**

Pour ce qui est de $p(]0, 1[)$, on écrira plutôt :
$$p(]0, 1[) = 1 - (p(0) + p(1)) = 1 - 0 = 1$$
On a donc un événement de probabilité 1, mais qui n'est pas l'univers !

Définition
Étant donné un espace probabilisé $(\Omega, \mathcal{T}, p)$,

- Si $A \in \mathcal{T}$ est tel que $A \neq \emptyset$ et $p(A) = 0$, on dit que A est *négligeable*.
- Si $A \in \mathcal{T}$ est tel que $A \neq \Omega$ et $p(A) = 1$, on dit que A est *presque sûr*.

I-3 Cohérence avec les propriétés connues

Remarque :

Les propriétés ci-dessous, déjà connues dans le cadre des probabilités finies se généralisent aisément à toutes les mesures de probabilités. (*à traiter en exercice*)

Propriété

On se donne un espace probabilisé $(\Omega, \mathcal{T}, p)$ et des événements $A, B \in \mathcal{T}$.

- Si A et B sont incompatibles, alors $p(A \sqcup B) = p(A) + p(B)$
- $p(\bar{A}) =$ 1 - p(A)
- Si $A \subset B$, on a $p(A) \leqslant p(B)$ et $p(B \backslash A) =$ p(B)-p(A).
- Pour A et B quelconques, $p(A \cup B)$ $= p(A) + p(B) - p(A \cap B)$.

Exercice 2

Montrer que
$p(A \cup B \cup C) = \Big(p(A) + p(B) + p(C)\Big) - \Big(p(A \cap B) + p(A \cap C) + p(B \cap C)\Big) + p(A \cap B \cap C)$

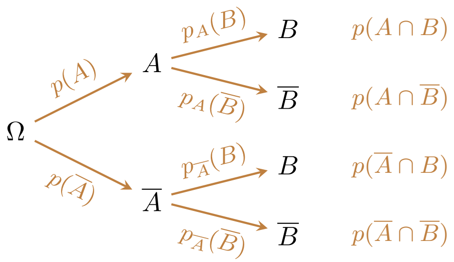
II Probabilités conditionnelles

II-1 Définition et exemples de calculs

Commentaires :

On rappelle que, dans le cas des probabilités finies, on sait que la probabilité d'un événement A connaissant déjà la réalisation d'un autre événement B tel que $p(B) \neq 0$ peut s'écrire (sous réserve de définition) $p_B(A) = \dfrac{p(A \cap B)}{p(B)}$

Cette formule peut s'interpréter grace de l'arbre de probabilité suivant :



Heureusement, on peut étendre cette notion à l'ensemble de toutes les mesures de probabilités.

Définition

Pour tout événement B de l'espace probabilisé $(\Omega, \mathcal{T}, p)$, vérifiant $p(B) \neq 0$ et $A \in \mathcal{T}$, on appelle *probabilité de A sachant B* et on note $p_B(A)$ la probabilité de l'événement A "sachant que B est déjà vérifié".

Propriété

Avec les notations de la définition précédente, l'application

$$p_B : \quad \mathcal{T} \quad \to \quad \mathbb{R}^+; \ A \mapsto P_B(A)$$

est bien une mesure de probabilités et on a

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

Démonstration : admise. ☐

Confusion

On voit bien sur la formule qu'il ne faut pas confondre la probabilité "conditionnelle" avec la probabilité de l'intersection !

Commentaires :

L'erreur fréquemment commise est de penser qu'il faut systématiquement utiliser la formule $P_B(A) = \frac{P(A \cap B)}{p(B)}$ pour calculer la probabilité conditionnelle, ce qui est faux. Voyons ci-dessous deux exemples de situation. Dans la première, on utilise la formule, la deuxième non.

Exercice 3

Un couple a deux enfants. Calculer la probabilité qu'ils aient 1 fille sachant qu'ils ont au moins un garçon.

Solution

En travaillant sur l'univers $\Omega = \{(\underbrace{i}_{1^{\text{er}} enfant}, \underbrace{j}_{2^{\text{ème}} enfant}) \in \{G, F\}^2\}$ avec la probabilité uniforme et en posant

A : « ils ont une fille » $\quad$ B : « ils ont au moins un garçon »

On a

$$p_B(A) = \frac{p(A \cap B)}{p(B)} = \frac{p(\{(F,G),(G,F)\})}{p(\{(F,G),(G,F),(G,G)\})} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}$$

Exercice 4

On dispose de 2 urnes appelées U_1 et U_2. L'urne U_2 contient autant de boules rouges que de boules bleues.
On lance maintenant un dé équilibré à 6 faces. On note RP : « le résultat est pair ».

Si le résultat est pair, on tire une boule dans l'urne U_2. On regarde la couleur de la boule obtenue. Décrire p_{RP} c'est-à-dire : donner l'univers, la tribu associée ainsi que la loi de probabilité p_{RP}.

Solution

On sait qu'on a tiré un nombre pair et qu'on tire alors une boule dans l'urne U_2 qui contient autant de boules rouges que bleues. Ainsi, on peut poser

$$\Omega = \{R, B\} \quad \text{où } R \text{ désigne le fait de tirer une boule rouge et } B \text{ une boule bleue.}$$

Alors Ω est fini. On peut donc poser comme tribu $\mathcal{T} = \mathcal{P}(\Omega)$. Étant donné qu'il y a le même nombre de boules rouges et bleues, la probabilité p_{RP} est alors l'équiprobabilité.

Commentaires :

En bref, lors d'expériences menées par étapes, il est souvent utile (et nécessaire) de déterminer la probabilité conditionnelle par d'autres moyens que la formule, par exemple par une analyse directe de la problématique en se plaçant directement dans le cadre du "sachant que". De plus, dans ce cadre, la probabilité conditionnelle sert justement à calculer les $P(A \cap B)$:

Exercice 5

Dans le cadre de l'exercice précédent, calculer la probabilité d'avoir un résultat pair pour le dé et d'obtenir une boule rouge.

Solution

On cherche $p(RP \cap \{R\})$:
Par probabilités conditionnelles, comme $p(RP) \neq 0$, on a

$$p(RP \cap \{R\}) = p(RP) p_{RP}(\{R\})$$

Or, on peut calculer par équiprobabilité sur les résultats du dé que

$$p(RP) = \frac{1}{2}$$

$$p_{RP}(\{R\}) = \frac{1}{2}$$

D'où

$$p(\text{« avoir un résultat pair pour le dé et obtenir une boule rouge »}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Propriété

Pour tout événement $A, B \in \mathcal{T}$ d'un espace probabilisé $(\Omega, \mathcal{T}, p)$, on a

$$p(A \cap B) = p_B(A)p(B) \quad \text{si } p(B) \neq 0 \quad \text{ou} \quad p(A \cap B) = p_A(B)p(A) \quad \text{si } p(A) \neq 0$$

Notation

Si $\boldsymbol{p(B) = 0}$, par convention, on va poser $\boldsymbol{p_B(A)p(B) = p(A \cap B) = 0}$.

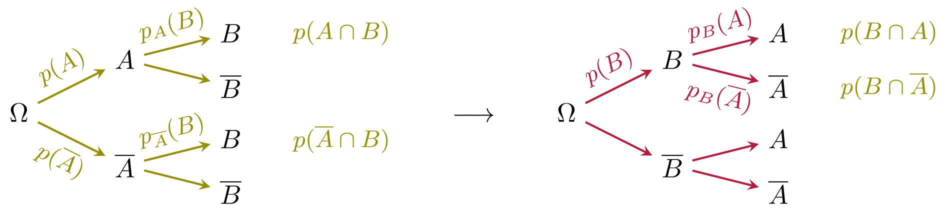
en effet, si $p(B) = 0$, comme $A \cap B \subset B$, on a également $p(A \cap B) = 0$. On généralise donc la formule $p_B(A)p(B) = p(A \cap B)$ dans ce cadre.
Les formules établies par la suite seront toutes écrites suivant cette convention.

II-2 Formules de Bayes

Après la généralisation de la formule des probabilités conditionnelles la formule de Bayes se généralise donc également sans modification :

Commentaires :

Rappelons que cette formule vise en quelque sorte à "inverser" les étapes, c'est-à-dire par exemple que si l'arbre de gauche ci-dessous est connu, on peut déterminer l'arbre de droite. (Cette formule n'a pas nécessité d'être apprise par coeur. Elle découle immédiatement des autres propriétés connues.)



Théorème *Formule de Bayes*

Soit $(\Omega, \mathcal{T}, p)$ un espace probabilisé. Pour tout $A, B \in \mathcal{T}$ tels que $p(A), p(B) \neq 0$, on a que

$$p_B(A) = \frac{p(A \cap B)}{p(B)} = \frac{p_A(B)p(A)}{p(B)} = \frac{p_A(B)p(A)}{p_A(B)p(A) + p_{\bar{A}}(B)p(\overline{A})}$$

II-3 Formule des probabilités composées

Théorème *Formule des probabilités composées*

Soient $(\Omega, \mathcal{T}, p)$ un espace probabilisé, n un entier supérieur ou égal à 2 et $A_1, \ldots, A_n \in \mathcal{T}$. On a alors

$$p(A_1 \cap \ldots \cap A_n) = p(A_1)p_{A_1}(A_2)p_{A_1 \cap A_2}(A_3)\ldots p_{A_1 \cap \ldots \cap A_{n-1}}(A_n)$$

Démonstration : exercice. □

Exercice 6

On dispose d'une urne contenant n boules dont 5 boules blanches. On tire sans remise 3 boules. Quelle est la probabilité d'obtenir 3 boules blanches ?

Solution

On note B_i l'événement "obtenir une boule blanche au $i^{\text{ème}}$ tirage". On cherche donc $p(B_1 \cap B_2 \cap B_3)$: D'après la formule des probabilités composées,

$$p(B_1 \cap B_2 \cap B_3) = p(B_1)p_{B_1}(B_2)p_{B_1 \cap B_2}p(B_3)$$

Or, par équirobabilité des tirages, on a

$$p(B_1) = \frac{5}{n} \quad ; \quad p_{B_1}(B_2) = \frac{4}{n-1} \quad ; \quad p_{B_1 \cap B_2}p(B_3) = \frac{3}{n-2}$$

D'où

$$p(B_1 \cap B_2 \cap B_3) = \frac{5}{n} \cdot \frac{4}{n-1} \cdot \frac{3}{n-2}$$

Commentaires :

Cette formule permet de démontrer des résultats que l'on considère souvent comme intuitifs, alors qu'ils ont en réalité (cf formule) une justification bien précise.

Remarque :

Cette formule est souvent utilisée quand les événements $A_1, \ldots, A_n$ sont donnés dans un ordre chronologique décroissant ; $A_n \subset A_{n-1} \subset \ldots \subset A_1$. La formule s'écrit alors

$$p(A_n) = p(A_1)p_{A_1}(A_2)\ldots p_{A_{n-1}}(A_n)$$

Exercice 7

On reprend l'exemple précédent. On cherche toujours la probabilité d'obtenir 3 boules blanches, mais avec la technique précédente.

Solution

On cherche donc $p(A_3)$ et on sait que $A_3 \subset A_2 \subset A_1$

D'après la formule des probabilités composées pour des événements chronologiques,

$$p(A_3) = p(A_1)p_{A_1}(A_2)p_{A_2}p(A_3)$$

Or, par équirobabilité des tirages, on a

$$p(A_1) = \frac{5}{n} \quad ; \quad p_{A_1}(A_2) = \frac{4}{n-1} \quad ; \quad p_{A_2}p(A_3) = \frac{3}{n-2}$$

D'où

$$p(A_3) = \frac{5}{n} \cdot \frac{4}{n-1} \cdot \frac{3}{n-2}$$

III Indépendance

On se donne un espace probabilisé $(\Omega, \mathcal{T}, p)$.

Commentaires :

Instinctivement un événement A est indépendant de B, si la réalisation de A ne dépend pas de celle de B. EN CONSEQUENCE : une restitution sous forme de formule serait donc la suivante :

$$p_B(A) = p(A)$$

D'un autre côté, cette formule a l'air non symétrique entre A et B (ce qui est contraire à l'instinct. Heureusement, ce n'est qu'une impression.

Propriété

On se donne $A, B \in \mathcal{T}$ tels que $p(A), p(B) \neq 0$. Alors on a

$$p_B(A) = p(A) \quad \Leftrightarrow \quad p(A \cap B) = p(A)p(B) \quad \Leftrightarrow \quad p_A(B) = p(B)$$

Démonstration :

Soient $A, B \in \mathcal{T}$ tels que $p(A), p(B) \neq 0$. Alors on a

$$p_B(A) = p(A) \Leftrightarrow p(A \cap B)/p(B) = p(A) \Leftrightarrow p(A \cap B) = p(A)p(B)$$

De même, on pourra écrire

$$p(A \cap B) = p(A)p(B) \Leftrightarrow p(B \cap A)/p(A) = p(B) \Leftrightarrow p_A(B)p(A)$$

□

Définition

Étant donné $A, B \in \mathcal{T}$, on dit que A et B sont indépendants si $p(A \cap B) = p(A)p(B)$.

Remarque :

Toutes les propriétés vue en première année sur l'indépendance restent vraies. *(Les relire et les redémontrer.)*

Remarque :

La formule d'indépendance est un équivalent de la question instinctive :

"les deux événements s'influencent-t-ils l'un l'autre ?"

Ainsi, il n'est pas forcément utile de se jeter immédiatement dans un calcul pour justifier ceci ! Un peu de bon sens suffit souvent à affirmer les choses ...

Par exemple : prenons A : "probabilité d'otenir Face sur la pièce équilibrée que je lance." et B : "probabilité qu'il fasse beau ce matin" n'ont clairement rien à voir l'un avec l'autre. On peut affirmer sans crainte qu'il y a indépendance...

Faux semblants

Attention toutefois aux affirmations trop hâtives ! Les apparences sont parfois trompeuses.

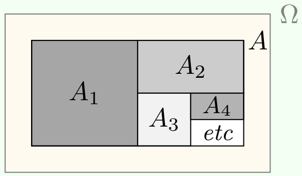
IV systèmes (quasi)-complets

IV-1 Définition

Définition

Étant donné $(\Omega, \mathcal{T}, p)$ un espace probabilisé, $A \in \mathcal{T}$ un événement et $I \subset \mathbb{N}$. On dit que la famille $(A_k)_{k\in I}$ est un *système complet d'événements* (ou une partition) de A si :

- $A_k \in \mathcal{T}$ pour tout $k \in I$;
- les événements A_k sont 2 à 2 incompatibles ;
- $\bigsqcup_{k\in I} A_k = A$



Commentaires :

Pour un ensemble A, la famille $(A_k)_{k\in I}$ dans $\mathcal{T}$ est donc un système complet d'événements de A si et seulement si :
- deux A_k distincts ne peuvent se produirent en même temps
- l'ensemble des A_k forme toutes les possibilités pour obtenir A.

Exemples :

10 ■ Si on cherche le rang d'apparition du premier pile lors d'une série de lancers de pile ou face, on pose

$$A_n \; : \; \text{"Obtenir Pile au rang } n\text{"} \quad \forall n \in \mathbb{N}^* \quad \text{et} \quad A_0 \; : \; \text{"Ne jamais obtenir Pile"}$$

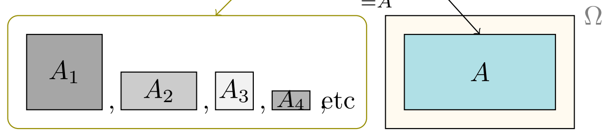
Alors $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements de l'univers.

11 ■ Si $\Omega = \mathbb{N}$ et A l'ensemble des nombres pairs. Alors, en posant $A_n = \{2n\}$ pour tout entier n, on trouve que $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements de A.

Notations

Faire attention aux mélanges de notations. Le système complet est une famille d'événements, pas une réunion ! Explication :

$$\{A_k\}_{k\in I} \neq \underbrace{\bigcup_{k\in I} A_k}_{=A}$$



Remarque :

On peut bien entendu avoir $A = \Omega$. On parle alors d'un système complet (ou partition) (de l'univers).

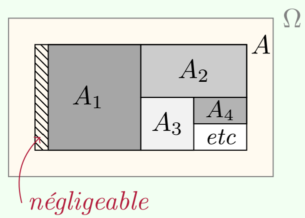
Commentaires :

Dans la suite (cas des variables discrètes ou à densité), nous auront besoin d'une notion plus souple, qui tient compte des événements négligeables, d'où la définition ci-dessous :

Définition

Étant donné $(\Omega, \mathcal{T}, p)$ un espace probabilisé, on dit que la famille $(A_k)_{k\in\mathbb{N}}$ est un *système quasi-complet d'événements* de $A \in \mathcal{T}$ si :

- $A_k \in \mathcal{T}$, $A_k \subset A$ pour tout $k \in \mathbb{N}$;
- les événements A_k sont 2 à 2 incompatibles ;
- $p\left(\bigsqcup_{k\in\mathbb{N}} A_k\right) = \sum_{k=0}^{+\infty} p(A_k) = p(A)$.



Convention d'écriture pour la suite du cours :
Dans la définition et dans la suite du cours, pour des facilités d'écriture, les indices de la réunion/somme dans les différents théorèmes et résultats sont pris dans $\mathbb{N}$ tout entier, mais rien n'empêche un changement d'indice, (par exemple une indexation sur $\mathbb{N}^*$ comme ci-dessous.)

Exemple 12 :

On imagine à nouveau qu'on lance une pièce de monnaie indéfiniment avec

$$\forall n \geqslant 1, \quad A_n = \text{« Obtenir Pile au } n^{\text{ième}} \text{ lancer »}$$

Montrons que $\{A_n\}_{n\in\mathbb{N}^*}$ **est quasi complet**. Les événements étant 2 à 2 incompatibles, il reste à montrer que $p\left(\bigcup_{n=1}^{+\infty} A_n\right) = 1$:

On sait par σ-additivité que

$$p\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} p(A_n)$$

Notons maintenant F_k : « obtenir Face au $n^{\text{ième}}$ lancer ». Alors pour $n \geqslant 1$:

$$A_n = F_1 \cap \ldots \cap F_{n-1} \cap \overline{F_n}$$

Les lancers étant infinis quoi qu'il arrive, les résultats à chaque lancer sont indépendants. Ainsi,

$$p(A_n) = p(F_1)\ldots p(F_{n-1})p(\overline{F_n})$$

Or, par équiprobabilité, au $i^{\text{ème}}$ lancer, on a une probabilité de $\frac{1}{2}$ d'obtenir Pile ou Face. Ainsi,

$$p(A_n) = \underbrace{\frac{1}{2}\cdot\ldots\cdot\frac{1}{2}}_{n-1 \text{ fois}}\cdot\frac{1}{2} = \left(\frac{1}{2}\right)^n$$

Maintenant :

$$p\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty}\left(\frac{1}{2}\right)^n = \frac{1}{1-\frac{1}{2}} - 1 = 1$$

Le système est donc bien quasi-complet...!

Remarque :

Dans l'exemple précédent, avec la notation A_0 : "Ne jamais obtenir Pile", comme $\{A_n\}_{n\in\mathbb{N}}$ est complet on peut en déduire que A_0 est négligeable :

$$p(A_0) = 1 - p\left(\bigcup_{n=1}^{+\infty} A_n\right) = 1 - 1 = 0$$

? Exercice 8

Réfléchir à ce qui change dans la démonstration de cet exemple si on suppose que l'on s'arrête dès l'obtention du premier Pile.

Solution

Les événements ne sont maintenant plus indépendants. En effet, si par exemple pile au premier lancer $(\overline{F_1})$, alors on ne peut pas obtenir face (F_2)au deuxième, puisque le jeu s'est arreté.... Il faut donc utiliser la formule des probabilités composées :

$$p(A_n) = p(F_1)p(F_1)_{F_1}\ldots p_{F_1\cap\ldots\cap F_{n-2}}(F_{n-1})p(\overline{F_n})$$

Mais fort heureusement, les probabilités sont les mêmes!!

Remarque :

Si dans la définition précédente, on prend $A_k = \emptyset$ pour tout $k > N$, on a en fait un système quasi-complet fini : $(A_k)_{0\leqslant k<N}$ où la somme $\sum_{k=0}^{+\infty} p(A_k)$ devient une somme finie $\sum_{k=0}^{N} p(A_k)$.

Exemple 13 :

On choisit un nombre au hasard entre 0 et 1 (0 et 1 inclu.).
On associe à cette expérience l'ensemble fondamental $\Omega = [0;1]$ et la tribu $\mathcal{B}$ des boréliens.

Comparons les systèmes suivants :

$$A_1 = [0;1/2[\text{ et } A_2 = [1/2;1] \qquad (1)$$

et

$$B_1 =]0;1/2[\text{ et } B_2 = [1/2;1] \qquad (2).$$

Le système (1) est clairement un système complet pour $\Omega = [0;1]$, car

$$A_1 \cap A_2 = \emptyset, \quad A_1 \cup A_2 = \Omega$$

Mais même si $B_1 \cap B_2 = \emptyset$,

$$B_1 \cup B_2 =]0;1] \neq \Omega$$

Donc le système (2) n'est **pas complet**. Néanmoins, la probabilité de tomber **exactement** sur 0 est nulle. (Pour l'instant, il faut se contenter de l'intuition, nous prouverons ceci un peu plus tard.) Ainsi, on a en fait

$$p(B_1 \cup B_2) = 1 - p(\overline{B_1 \cup B_2}) = 1 - p(\{0\}) = 1$$

Le système (2) est donc quasi-complet.

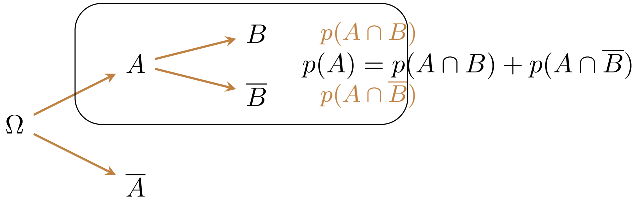
Commentaires :

En observant les définitions, on voit qu'un système complet est un cas particulier de système quasi-complet, toutes les propriétés suivantes sont donc valables dans les deux cadres...

IV-2 Formule(s) des probabilités totales

Commentaires :

Si la formule des probabilités composées consiste à suivre **une branche** *d'un arbre de probabilités, la formule des probabilités totales ci-dessous consistera plutôt à analyser la décomposition suivant* **ses extrémités.**



Voyons par étape comment ceci se généralise à des systèmes infinis :

Propriété *(Découpage par système quasi-complet)*

On se donne un espace probabilisé $(\Omega,\mathcal{T},p)$. Soit B un événement quelconque. Si $(A_n)_{n\in\mathbb{N}}$ est un **système (quasi)-complet d'événements** de Ω (ou de A tel que $B\subset A$), alors

$$p(B)=p(B\cap \bigsqcup_{n\in\mathbb{N}} A_n)=\sum_{n=0}^{+\infty} p(B\cap A_n).$$

Démonstration :

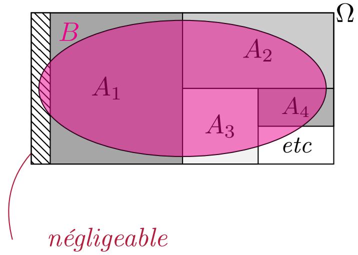
Posons C le complémentaire dans Ω de $\bigsqcup_{n=0}^{+\infty} A_n$. Alors $\Omega = C\sqcup \bigsqcup_{n=0}^{+\infty} A_n$ puis

$$\begin{aligned} p(B) &= p(B\cap\Omega)=p(B\cap C\sqcup \bigsqcup_{n=0}^{+\infty} A_n)\\ &= p\big((B\cap C)\sqcup(B\cap \bigsqcup_{n=0}^{+\infty} A_n)\big)=p(B\cap C)+P(B\cap \bigsqcup_{n=0}^{+\infty} A_n)\end{aligned}$$

Or, $B\cap C\subset C$, d'où $0\leq P(B\cap C)\leq P(C)=0$ et donc $P(B\cap C)=0$.
On en déduit que

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

□



? Exercice 9

On lance une pièce de monnaie indéfiniement. Calculer la probabilité de l'événement B :« le deuxième Pile arrive juste après le premier. »

Solution

Nous avons démontré tout à l'heure qu'en posant $A_n=$ « Obtenir Pile au n$^{\text{ième}}$ lancer » pour tout $n\in\mathbb{N}^*$, le système $\{A_n\}_{n\in\mathbb{N}^*}$ est quasi complet. Ainsi, par découpage en système quasi-complet, on a

$$P(B)=\sum_{n=1}^{+\infty} P(B\cap A_n)$$

Or, en reprenant les notations utilisées dans la démonstration de "$\{A_n\}_{n\in\mathbb{N}^*}$ est quasi complet" à la page 9, on trouve que

$$B\cap A_n = F_1\cap\ldots\cap F_{n-1}\cap\overline{F_n}\cap\overline{F_{n+1}}$$

D'où la encore par indépendance,

$$p(B\cap A_n)=p(F_1)\ldots p(F_{n-1})p(\overline{F_n})p(\overline{F_{n+1}})=\left(\frac{1}{2}\right)^{n+1}$$

Et donc

$$p(B)=\sum_{n=1}^{+\infty}\left(\frac{1}{2}\right)^{n+1}=\left(\frac{1}{2}\right)^{2}\sum_{n=1}^{+\infty}\left(\frac{1}{2}\right)^{n-1}=\left(\frac{1}{2}\right)^{2}\frac{1}{1-\frac{1}{2}}=\frac{1}{2}$$

(intuitif et logique. Après un pile arrive un autre pile avec « 1 chance sur 2 » !)

Théorème (Formule des probabilités totales)

Étant donné $(\Omega,\mathcal{T},p)$ un espace probabilisé, et $(A_0,\ldots A_n,\ldots)$ un système quasi-complet d'événements de $(\Omega,\mathcal{T},p)$. Alors, pour tout événement $B\in\mathcal{T}$, on a

$$p(B)=\sum_{k=0}^{+\infty} p_{A_k}(B)p(A_k)\qquad \textit{(avec } p_{A_k}(B)p(A_k)=0 \textit{ si } p(A_k)=0)$$

`Démonstration :`

On sait déjà, d'après le découpage en système quasi-complet démontré précédemment, que

$$p(B)=\sum_{k=0}^{+\infty} p(B\cap A_k) \qquad (*)$$

Or, pour tout $k=0,\dots,n$, on a $p(B\cap A_k)=p_{A_k}(B)p(A_k)$. On réinjecte dans $(*)$ et c'est terminé. □

■ Exemple 14 :

Suivant une étude (non contractuelle), on sait que la probabilité p_k pour une famille d'avoir k enfants est telle que

$$p_0=p_1=a, \quad p_k=(1-2a)2^{-(k-1)}\ \forall k\geqslant 2.$$

On part du principe que la probabilité d'avoir une fille ou un garçon est la même. Quelle est dans ce cas la probabilité d'avoir exactement 2 garçons sur l'ensemble des enfants eu par le couple ?

Notons G l'événement "avoir deux garçons" et E_n : "La famille a n enfants." Alors le système $\{E_n\}_{n\geqslant 2}$ est un système complet de l'événement G et on a alors

$$P(G)=\sum_{n=2}^{+\infty}P(G\cap E_n)=\sum_{n=2}^{+\infty}P_{E_n}(G)P(E_n)$$

Or $P_{E_n}(G)=P(\text{"Il y a deux garçons et } n-2 \text{ filles"})$ L'événement ci-dessus est l'ensemble des permutations de la liste $GG\underbrace{F\dots F}_{n-2 \text{ fois}}$.

De plus, chacune de ces listes est de probabilité égale à $\frac{1}{2^n}$ par indépendance. Ainsi, comme on a $\binom{n}{2}$ permutations possibles,

D'où
$$P_{E_n}(G)=\binom{n}{2}\frac{1}{2^n}=\frac{n(n-1)}{2}\frac{1}{2^n}$$

$$P(G)=\sum_{n=2}^{+\infty}\frac{n(n-1)}{2}\frac{1}{2^n}(1-2a)2^{-(n-1)}=(1-2a)\sum_{n=2}^{+\infty}n(n-1)\frac{1}{4^n}=(1-2a)\frac{1}{4^2}\frac{2}{\left(1-\frac{1}{4}\right)^3}$$

i.e.
$$P(G)=\frac{8}{27}(1-2a)$$

V Variables aléatoires

Rappel : Dans le cadre des variables aléatoires **finies**, en première année, avec un univers Ω fini, nous avons défini une variable aléatoire simplement comme une application

$$X:\Omega\to\mathbb{R}.$$

Nous avons alors régulièrement utilisé des notations de type « $X=i$ » pour désigner l'événements "la variable aléatoire X prend la valeur i" (pour $i\in\mathbb{R}$), ou « $X\in B$ » pour désigner l'événement "la variable aléatoire X prend ses valeurs dans B" (pour B un ensemble de réels).

Rappelons que cela signifie la chose suivante :

$$\boldsymbol{(X=i)}\ =\{w\in\Omega\mid X(w)=i\} \qquad \textit{(antécédents de } i \textit{ par } X\textit{)}$$

$$\boldsymbol{(X\in B)}\ =\{w\in\Omega\mid X(w)\in B\} \qquad \textit{(antécédents de tous les éléments de } B \textit{ par } X\textit{)}$$

■ Exemple 15 :

On lance deux dés et on appelle S la somme obtenue. Pour l'univers

$$\Omega=\{(i,j)\mid i,j\in[\![1;6]\!]\}$$

et

$$S:\Omega\to\mathbb{R};\ (i,j)\mapsto i+j$$

l'événement « $S=4$ » correspond à

$$(S=4)=\{w\in\Omega\mid S(w)=4\}=\{(1,3);(2;2);(3,1)\}$$

De même « $S\leqslant 3$ » correspond à

$$(S\leqslant 3)=\{w\in\Omega\mid S(w)\leqslant 3\}=\{(1,1);(1;2);(2,1)\}$$

et à y regarder de plus près, « $S\in]2,3[$ » correspond à

$$(S\in]2,3[)=\emptyset$$

V-1 Variables aléatoires quelconques

Commentaires :

Dans le cadre futur des variables construites sur un univers étant un intervalle I de $\mathbb{R}$, il faut adapter la théorie à la présence d'une tribu potentiellement différente de $\mathcal{P}(\Omega)$:

Définition

On appelle variable aléatoire réelle sur $(\Omega,\mathcal{T},P)$ toute fonction $X:\Omega\to\mathbb{R}$ telle que :

1. $(\Omega,\mathcal{T},P)$ soit un espace probabilisé
2. pour tout $a\in\mathbb{R}$, $(X\leq a)\in\mathcal{T}$.

Remarque :

La deuxième condition peut également se traduire ainsi :

$$\text{pour tout } a \in \mathbb{R},\ (X \leq a) \text{ est un événement.}$$

Il se trouve que grâce aux propriétés des tribus, tous les ensembles de type « $X \in B$ » avec B raisonnable auront également un sens et seront également des événements bien définis :

Propriété

Si X est une variable aléatoire et I un intervalle ou une réunion finie d'intervalles, alors $(X \in I)$ est un événement.

Commentaires :

Le but n'étant pas de rentrer dans des excès de technicité, on ne fera pas d'exemple chiffré ici. On se contentera des faits suivants, qui nous assurent que grand nombres de fonctions sont des variables aléatoires. (Vous n'êtes pas censés savoir démontrer qu'une fonction est une variable aléatoire.)

■ **Exemple 16** :

Pour Ω fini, toute fonction $X : \Omega \to \mathbb{R}$ est une variable aléatoire sur $(\Omega, \mathcal{P}(\Omega), P)$ quelquesoit la probabilité p.

Proposition

Si Ω est un intervalle ou une réunion finie d'intervalle de $\mathbb{R}$, toute fonction continue ou continue par morceaux est une variable aléatoire sur $(\Omega, \mathcal{B}(\Omega), P)$.

Proposition

Soient X, Y deux variables aléatoires réelles définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$, alors

1. $X + Y$ est une variable aléatoire.
2. XY est une va.
3. Si Y ne s'annule pas, $\dfrac{X}{Y}$ est une variable aléatoire.

Proposition

Soit f une fonction continue sur $\mathcal{D}_f \subset \mathbb{R}$ et X une variable aléatoire sur $(\Omega, \mathcal{T}, P)$ telle que $X(\Omega) \subset \mathcal{D}_f$, alors $f \circ X$ est une variable aléatoire.

■ **Exemple 17** :

Si X, Y, Z sont des va sur $(\Omega, \mathcal{T}, P)$, alors $\ln \frac{|XY-Z|+1}{1+Z^2}$ est une va.

V-2 Loi d'une v.a. par fonction de répartition

V.2-a) Valeurs de la variable : notion de support

Commentaires :

L'intérêt d'une variable aléatoire sur $(\Omega, \mathcal{T}, P)$ est notamment de pouvoir "se débarraser" de l'univers d'origine, sur lequel il était peut être difficile d'effectuer des calculs.
On constate qu'en pratique, il n'y a nul besoin de connaître Ω. En revanche, la probabilité P est bien celle de l'espace $(\Omega, \mathcal{T}, P)$. En général, on ne sait pas la décrire intégralement, mais on souhaite pouvoir calculer les probabilités du type $P(X \in B)$ ou $P(X = b)$ ou encore $P(X \leqslant b)$.

■ **Exemple 18** :

Effectuons une infinité de lancers d'une pièce équilibrée. L'univers que l'on peut associer à cette expérience est le suivant :

$$\Omega = \{(u_n)_{n \in \mathbb{N}^*} \mid u_n \in \{P, F\}\}$$

Un élément est donc une suite (infinie) de F et de P qui donne l'ensemble des faces obtenues chaque lancer, par exemple

$$(F, F, P, F, P, \dots)$$

La probabilité d'un tel événement n'est pas calculable à l'aide de l'indépendance ou des probabilités composées (car on a une infinité de valeurs...)... *(à méditer)*
En revanche, si on note maintenant

$$X : \Omega \to \mathbb{N};\ w \mapsto \begin{cases} \text{le rang d'apparition du premier Pile s'il apparait au moins une fois} \\ 0 \text{ sinon} \end{cases}$$

X sera bien une variable aléatoire (admis). Si on ne s'intéresse maintenant qu'à la seule valeur de X, on peut facilement calculer $P(X = n)$ pour tout $n \in \mathbb{N}$. On l'a déjà fait dans l'exemple sur le système complet p.9 :

$$P(X = n) = \begin{cases} \frac{1}{2^n} & \forall n \in \mathbb{N}^* \\ 0 & \text{si } n = 0 \end{cases}$$

Commentaires :

*Dans l'exemple ci-dessus, X a d'un côté un **univers de définition** constitué des suites $(P, F, \dots)$ et de l'autre un ensemble de **valeurs** $0, 1, \dots$.*
De manière générale, pour toute fonction, l'ensemble de ses valeurs est défini par :

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

Notons alors que, comme il s'agit de l'ensemble de toutes les valeurs possibles, on a

$$\boldsymbol{P(X \in X(\Omega)) = P(\Omega) = 1.}$$

■ Exemples :

19 ■ Sur l'exemple précédent (lancer infini d'une pièce et X rang d'appatition du premier Pile), on a

$$X(\Omega) = \mathbb{N}$$

20 ■ Si X est le résultat du choix d'un nombre dans l'intervalle $[0, 1]$ alors trivialement

$$X(\Omega) = [0, 1]$$

21 ■ Si on regarde le nombre de boules rouges que l'on peut obtenir en sélectionnant de manière simultanée 10 boules dans une urne contenant 6 boules rouges et 6 boules blanches, on a

$$X(\Omega) = [\![4, 6]\!]$$

Commentaires :

Étant donné la possibilité d'avoir des événements négligeables, on aimerait se laisser un peu plus de souplesse. On aimerait par exemple, dans l'exemple de la permière apparition de Pile, pouvoir se débarasser du cas $X = 0$ qui est négligeable :

Définition

Si X est une v.a. définie sur l'espace probabilisé $(\Omega, \mathcal{T}, P)$, on appelle *support de X* et on note $Supp(X)$ tout sous-ensemble de $X(\Omega)$ tel que $P\big(X \in Supp(X)\big) = 1$.

Commentaires :

Dit "simplement" cela correspond à l'ensemble des valeurs que peut prendre la variable X où on se se préoccuperait pas de ce qui est négligeable. On va donc voir qu'un support n'est pas forcément défini de manière unique (cf ci-dessous)

■ Exemples :

22 ■ Un support trivial (**et maximum !**) pour toute variable aléatoire est $X(\Omega)$:

En effet, on vient de revoir plus haut que

$$P(X \in X(\Omega)) = 1.$$

23 ■ Sur la série infinie de lancers d'une pièce, si X est le rang d'apparition du premier Pile, comme $P(X \in \mathbb{N}^*) = 1$, on peut poser

$$Supp(X) = \mathbb{N}^*$$

24 ■ Si X est le résultat du choix d'un nombre dans l'intervalle $[0, 1]$ alors (sous-réserve d'admettre pour l'instant que 0 est de probabilité nulle), on a par exemple

$$Supp(X) =]0, 1]$$

mais il y a d'autres possibilités ! En effet, si 0 est de probabilité nulle, il en va certainement de même pour 1, ainsi

$$Supp(X) =]0, 1[$$

Et on pourrait peut être en retirer encore d'autres (mais tout ceci arrivera un peu plus tard…)

25 ■ Si on regarde le nombre de boules rouges que l'on peut obtenir en sélectionnant de manière simultanée 10 boules dans une urne contenant 6 boules rouges et 6 boules blanches, aucun élément de $[\![4, 6]\!]$ n'est de probabilité nulle. Ainsi, on ne peut rien retirer et on ne peut trouver de support plus petit que

$$Supp(X) = [\![4, 6]\!]$$

Notation tolérée

Souvent, dans les ouvrages (et certaines fois au concours), les notions $X(\Omega)$ et $Supp(X)$ sont confondues. C'est un abus de notation qui est toléré. Néanmoins, dans ce cours, la différence sera faite systématiquement afin de ne pas commettre d'erreur de réflexion.

V.2-b) Fonction de répartition d'une variable aléatoire

Commentaires :

On a vu qu'une variable aléaoire ne dépendait pas de l'univers sur lequel elle est constuite mais seulement de son support. Pour déterminer sa loi, dans tous les cas, on peut utiliser la "fonction de répartition"; La notion de "loi" d'une variable aléatoire commune à toutes les variables est celle de la fonction de répartition, qui, on va le voir, permet de décrire complètement l'ensemble des probabilités d'une variable quelle qu'elle soit.

Définition

Étant donnée une variable aléatoire X, on appelle *fonction de répartition* de X la fonction

$$F_X : \begin{array}{rcl} \mathbb{R} & \to & [0,1] \\ t & \mapsto & P(X \leqslant t) \end{array}$$

■ Exemple 26 :

On choisit simultanément 10 boules dans une urne contenant 6 boules rouges et 6 boules blanches. On note X le nombre de boules rouges obtenues et on note F la fonction de répartition de X. On cherche à déterminer $F(2), F(4), F(4.5), F(5), F(6)$

⋆ On a facilement que $F(2) = P(X \leqslant 2) = 0$ car l'événement est impossible.

⋆ A contrario, $F(6) = P(X \leqslant 6) = 1$ car c'est l'événement certain.

⋆ On a $F(4) = P(X \leqslant 4) = P(X = 4)$ car il n'y a pas d'autre possibilité.

⋆ Comme les valeurs de X sont nécessairement entières, on a $F(4,5) = P(X \leqslant 4.5) = P(X \leqslant 4) = F(4)$.

⋆ Pour la même raison, on a $F(5) = P(X \leqslant 5) = P(X = 4) + P(X = 5)$.

Pour résumer après avoir fini les calculs (exercice) :

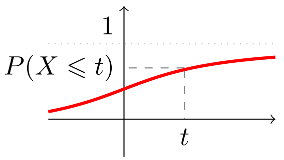
k	2	4	4.5	5	6
$P(X = k)$	0	$\frac{5}{22}$	$\frac{5}{22}$	$\frac{6}{11}$	$\frac{5}{22}$

Remarque :

Dans la situation de l'exemple précédent, on a trouvé les valeurs de F seulement pour $t \in \{2,\ 4,\ 4.5,\ 5,\ 6\}$. Elle n'est donc pas entièrement déterminée. Il faudrait préciser l'ensemble de tous les $F(t)$ pour tous les t réels. On en est loin !

Commentaires :

La fonction de répartition est une fonction réelle tout à fait classique, qui n'est **pas** *une variable aléatoire. Elle peut se dessiner à l'aide d'un graphique tout à fait habituel. Elle peut par exemple avoir cette allure là, (avec toujours des ordonnées clairement comprises entre 0 et 1) :*



En abscisse : les valeurs de t **pour tout t réel** *et en ordonnées, les probabilités associées. Nous allons commencer par voir quelques valeurs particulières (que nous aurons commencer à aborder dans l'exemple précédent.)*

Propriété

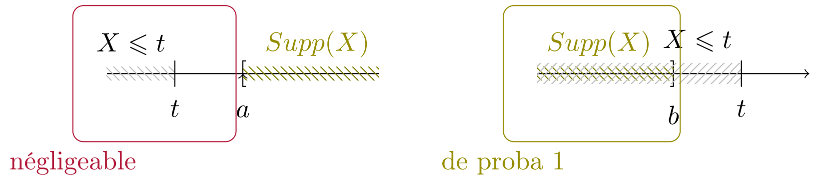
Si $Supp(X) \subset [a, +\infty[$, alors $F_X(t) = 0 \quad \forall t < a$.
De même, si $Supp(X) \subset]-\infty, b]$, alors $F_X(t) = 1 \quad \forall t \geqslant b$

`Et en français ?? :`

⋆ *Si les valeurs de X sont toutes supérieurs à a, alors, si $t < a$, l'événement $X \leqslant t$ est impossible.*

⋆ *Si les valeurs de X sont toutes inférieures à b, alors, si $t \geqslant b$, l'événement $X \leqslant b$ est certain.*

`Illustrations :`



`Démonstration :`

• Si $Supp(X) \subset [a, +\infty[$:

Soit $t < a$. Alors $(X \leqslant t) \subset (X < a) \subset (X \in \overline{Supp(X)})$. D'où

$$0 \leqslant P(X \leqslant t) \leqslant P(X < a) \leqslant P(X \in \overline{Supp(X)}) = 1 - P(X \in Supp(X)) = 1 - 1 = 0$$

et donc $P(X \leqslant t) = 0$.

• Si $Supp(X) \subset]-\infty, b]$:

Soit $t \geqslant b$. Alors $(X \in Supp(X)) \subset (X \leqslant b) \subset (X \leqslant t)$. D'où

$$1 = P(X \in Supp(X)) \leqslant P(X \leqslant b) \leqslant P(X \leqslant t) \leqslant 1$$

et donc $P(X \leqslant t) = 1$.

□

De plus, il vient rapidement une propriété élémentaire :

Propriété

Soit F_X la fonction de répartition d'une variable aléatoire X. Alors F_X est une fonction croissante.

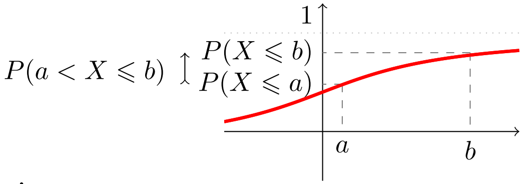
`Démonstration :`

Soient $u, v \in \mathbb{R}$ tels que $u \leqslant v$. Comme $(X \leqslant u) \subset (X \leqslant v)$, on obtient immédiatement $F_X(u) = P(X \leqslant u) \leqslant P(X \leqslant v) = F_X(v)$. □

Continuons l'interprétation :

Propriété

Si $a < b$, $P(a < X \leqslant b)$ désigne la "différence de hauteur" entre $P(X \leqslant a)$ et $P(X \leqslant b)$.

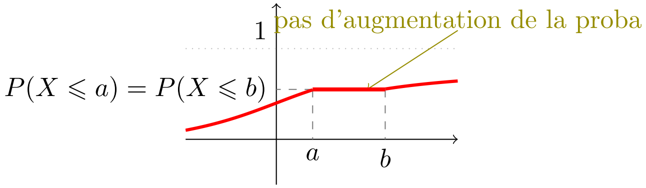


Démonstration :
$P(a < X \leqslant b) = P(X \leqslant b \setminus X \leqslant a) = P(X \leqslant b) - P(X \leqslant a)$ □

De la croissance de F_X et du résultat précédent, on tire également que :

Propriété

Soient $a, b \in \mathbb{R}$ tels que $a < b$. Alors
$$P(a < X \leqslant b) = 0 \text{ ssi } F_X \text{ est constante sur } [a, b[.$$



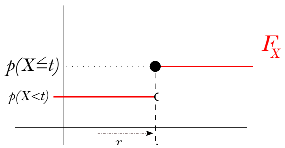
Voyons maintenant les inégalités strictes :

Propriété

Soit F_X la fonction de répartition d'une v.a. X et $t \in \mathbb{R}$. Alors
$$P(X < t) = \lim_{\substack{x \to t \\ x < t}} P(X \leqslant x) = \lim_{\substack{x \to t \\ x < t}} F_X(x)$$

Démonstration : admise.
□

Explication sur un graphique :



Remarque :

Soit F_X la fonction de répartition d'une v.a. X. Alors,
$$P(X = t) = P(X \leqslant t) - P(X < t)$$
Ce qui veut dire que :
$P(X = t) \neq 0$ si et seulement si t est un point de discontinuité de F_X
et
$P(X = t)$ est la "hauteur" de la discontinuité de la fonction de répartition en t.

Commentaires :

Nous allons voir maintenant 2 exemples : un dans le cadre des probabilités finies, et l'autre dans le cadre d'un support $[0, 1]$

■ Exemple 27 :

Revenons à : "On choisit simultanément 10 boules dans une urne contenant 6 boules rouges et 6 boules blanches. On note X le nombre de boules rouges obtenues." On cherche à établir la fonction de répartition F_X de X.

- Tout d'abord, le support :

C'est une variable finie avec $Supp(X) = \{4, 5, 6\} \subset [4, 6]$. Ainsi,
$$F_X(x) = \begin{cases} 0 & \text{si } x < 4 \\ 1 & \text{si } x \geqslant 6 \end{cases}$$

- Les intervalles où F_X est constante :

On sait que $P(4 < X < 5) = P(5 < X < 6) = 0$, ainsi, F_X est constante entre $4, 5$ puis $5, 6$.

- Les valeurs restantes :

Nous avons déjà déterminé les valeurs en $4, 5, 6$ dans un exemple précédent. On avait

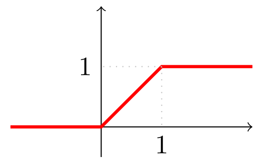
k	4	5	6
$P(X = k)$	$\frac{5}{22}$	$\frac{6}{11}$	$\frac{5}{22}$

D'où la fonction $F_X(x) = \begin{cases} 0 & \text{si } x < 4 \\ \frac{5}{22} & \text{si } 4 \leqslant x < 5 \\ \frac{5}{22} + \frac{6}{11} & \text{si } 5 \leqslant x < 6 \\ 1 & \text{si } x \geqslant 6 \end{cases}$

■ Exemple 28 :

Loi uniforme sur $]0; 1]$

On choisit un nombre X au hasard dans l'intervalle $]0; 1]$. Montrons que la fonction de répartition de X est F_X définie comme suit :
$$F_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ x & \text{si } x \in]0; 1[\\ 1 & \text{si } x \geqslant 1 \end{cases}$$



- Tout d'abord, $Supp(X) =]0; 1]$, d'où :
$$F_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 & \text{si } x \geqslant 1 \end{cases}$$

- Soit $x \in]0; 1[$. Montrons que $F_X(x) = x$:

De manière intuitive, on peut dire que la probabilité d'obtenir un nombre dans l'intervalle $]0;x]$ est x fois la probabilité d'obtenir un nombre dans $]0;1]$. i.e.

$$P(0 < X \leqslant x) = x.P(0 < X \leqslant 1) = x$$

Nous allons le montrer en plusieurs étapes : d'abord $x = \frac{1}{n}$ pour $n \in \mathbb{N}^*$, puis $x = \frac{k}{n} \in]0,1]$ pour $n, k \in \mathbb{N}^*$, et enfin x réel quelconque dans $]0,1]$. Pour commencer, remarquons qu'on a

$$F_X(x) = P(X \leqslant x) = P(0 < X \leqslant x)$$

$\star$ Pour $x = \frac{1}{n}$ avec $n \in \mathbb{N}^*$:

Par choix équiprobable, pour tout $n \in \mathbb{N}^*$, on a

$$P\left(0 < X \leqslant \frac{1}{n}\right) = P\left(\frac{1}{n} < X \leqslant \frac{2}{n}\right) = \ldots = P\left(\frac{n-1}{n} < X \leqslant \frac{1}{n}\right)$$

Sommons ces valeurs. On obtient d'une part

$$\sum_{k=0}^{n-1} P\left(\frac{k}{n} < X \leqslant \frac{k+1}{n}\right) = P(]0;1]) = 1$$

et d'autre part

$$\sum_{k=0}^{n-1} P\left(\frac{k}{n} < X \leqslant \frac{k+1}{n}\right) = n.P\left(0 < X \leqslant \frac{1}{n}\right)$$

d'où

$$p\left(0 < X \leqslant \frac{1}{n}\right) = \frac{1}{n}$$

$\star$ Pour $x = \frac{k}{n}$ avec $n, k \in \mathbb{N}^*$ où $k \leqslant n$:

Comme $\left\{\left(\frac{j}{n} < X \leqslant \frac{j+1}{n}\right)\right\}_{j=0\ldots,k-1}$ et un système complet de $\left(0 < X \leqslant \frac{k}{n}\right)$, on a

$$P\left(0 < X \leqslant \frac{k}{n}\right) = \sum_{j=0}^{k-1} P\left(\frac{j}{n} < X \leqslant \frac{j+1}{n}\right) = \frac{k}{n}$$

$\star$ Pour x réel quelconque dans $]0,1]$:

On cherche à encadrer x entre deux nombres rationnels du type $\frac{k}{n}$ et $\frac{k+1}{n}$. Remarquons alors que pour tout $x \in]0;1[$ et tout $n \in \mathbb{N}^*$, on a

$$\frac{k}{n} \leqslant x < \frac{k+1}{n} \quad \Leftrightarrow \quad k \leqslant nx < k+1 \quad \Leftrightarrow \quad k = \lfloor nx \rfloor$$

D'où, si $x \in]0,1[$, en posant $k = \lfloor nx \rfloor$, par croissance de p, pour tout $n \in \mathbb{N}^*$, on a

$$\underbrace{P(0 < X \leqslant \frac{k}{n})}_{=\frac{k}{n}} \leqslant P(0 < X \leqslant x) \leqslant P(0 < X < \frac{k+1}{n}) \leqslant \underbrace{P(0 < X \leqslant \frac{k+1}{n})}_{=\frac{k+1}{n}}$$

D'où

$$\frac{\lfloor nx \rfloor}{n} \leqslant P(0 < X \leqslant x) \leqslant \frac{\lfloor nx \rfloor + 1}{n}$$

et, par passage à la limite quand n tend vers l'infini, et théorème des gendarmes $\lim\limits_{n\to+\infty} \frac{\lfloor nx \rfloor}{n} = \lim\limits_{n\to+\infty} \frac{\lfloor nx \rfloor+1}{n} = x$ (exercice en considérant que $\lfloor nx \rfloor \in [nx-1, nx]$)

$$P(0 < X \leqslant x) = x$$

ce qui achève la démonstration.

Au final, au constate sur une fonction de répartition que :

Propriété

Soit F_X la fonction de répartition d'une variable aléatoire X. Alors, les conditions suivantes sont vérifiées :

i) $0 \leqslant F_X(t) \leqslant 1 \qquad \forall t \in \mathbb{R}$.

ii) F_X est une fonction croissante.

iii) F_X est une fonction continue à droite en tout point $t \in \mathbb{R}$.

iv) $\lim\limits_{t\to-\infty} F_X(t) = 0$ et $\lim\limits_{t\to+\infty} F_X(t) = 1$

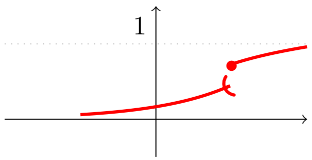
Démonstration :

i) Soit $t \in \mathbb{R}$. $F_X(t) = P(X \leqslant t) \in [0;1]$. ii) déjà vu

le reste est hors programme. □

Commentaires :

La condition iii) prendra tout son sens dans le futur chapitre sur les variables aléatoires discrètes. Elle consiste à dire qu'en cas de discontinuité, les points inclus se présentent toujours le la manière suivante :



V.2-c) Fonction dite "de répartition"

Définition

De manière générale, toute fonction $F : \mathbb{R} \to \mathbb{R}$ vérifiant les conditions

i) $0 \leqslant F(t) \leqslant 1 \quad \forall t \in \mathbb{R}$.

ii) F est une fonction croissante.

iii) F est une fonction continue à droite en tout point $t \in \mathbb{R}$.

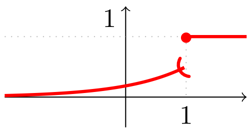
iv) $\lim_{t\to-\infty} F(t) = 0$ et $\lim_{t\to+\infty} F(t) = 1$

est appelée *fonction de répartition.*

Commentaires :

Cette fonction peut être définie de manière indépendante d'une quelconque variable aléatoire. Si on définit par exemple la fonction suivante :

$$F(x) = \begin{cases} \frac{e^x}{2e} & si\ x < 1 \\ 1 & si\ x \geqslant 1 \end{cases}$$



Alors F vérifie bien les conditions ci-dessus. C'est une fonction de répartition, et elle n'a pas été construite à partir d'une expérience. Néanmoins, les variables aléatoires ne sont jamais bien loin :

Propriété

Pour toute fonction de répartition F, il existe une variable aléatoire X de fonction de répartition F_X telle que

$$F = F_X$$

Autrement dit, toute à toute fonction de répartition, on peut associer une variable aléatoire.

Démonstration : admise. □

Commentaires :

Avec un peu plus d'habitude, nous verrons que nous pouvons déduire des fonctions de répartition toutes les informations utiles : support, type de variable, différentes probabilités, etc...Nous verrons ceci plus en détail dans les chapitres spécifiques.

V-3 Indépendance de variables aléatoires

V.3-a) Cas de deux variables

Définition

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$. On dit que X et Y sont *indépendantes* si, pour tout $A, B \in \mathcal{T}$, on a

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

Remarque :

Tout comme dans le cas des événements, la définition est une traduction mathématique de l'idée instinctive :

Le résultat d'une des variables n'influence pas l'autre.

Il est donc inutile de se jeter dans des calculs (qui peuvent être fastidieux !) quand la situation est claire.

Proposition

Deux variables aléatoires X, Y définies sur un même espace probabilisé sont indépendantes si et seulement si, pour tout $x, y \in \mathbb{R}$,

$$P(X \leqslant x \cap Y \leqslant y) = F(x)F(y)$$

Démonstration : admise. □

Exercice 10

On choisit deux nombres au hasard de manière indépendante dans l'intervalle $]0; 1]$. Quelle est la probabilité d'obtenir deux nombres inférieurs ou égaux à $\frac{1}{2}$?

Solution

On pose X et Y les variables donnant les deux résultats. Ce sont deux variables aléatoires indépendantes suivant des lois uniformes sur $]0; 1]$.Alors

$$P(X \leqslant \frac{1}{2}, Y \leqslant \frac{1}{2}) = F_X\left(\frac{1}{2}\right) F_Y\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

V.3-b) De manière générale :

Définition

Soient $(X_k)_{k\in K}$ une famille (finie ou infinie) de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$. On dit que les variables X_k sont *mutuellement indépendantes* si, pour toute famille $(A_k)_{k\in\mathbb{N}}$ d'événements, on a

$$P\left(\bigcap_{k\in K}(X_k \in A_k)\right) = \prod_{k\in K} P\left(X_k \in A_k\right)$$

et voici quelques propriétés de bon sens :

Propriété

Si $(X_k)_{k\in K}$ une famille de variables aléatoires mutuellement indépendantes, alors toute sous-famille de variables est mutuellement indépendante.

Démonstration : exercice. □

Lemme *des coalitions*

Si $(X_1, \ldots, X_n, X_{n+1}, \ldots X_{n+p})$ est une famille de variables mutuellement indépendantes, alors, pour tout fonction, $u : \mathbb{R}^n \to \mathbb{R}$, $v : \mathbb{R}^p \to \mathbb{R}$, les variables $u(X_1, \ldots, X_n)$ et $v(X_{n+1}, \ldots, X_{n+p})$ sont indépendantes.

Commentaires :

Le lemme ci-dessus signifie simplement qu'on peut associer les variables indépendantes de toute manière souhaitée. Pour peu qu'on les divise au moins en 2 paquets disjoints, ça reste bien indépendant.