Espaces probabilisés.

Chapitre 06 : cours complet.

1. Ensembles dénombrables.

Définition 1.1 : ensemble fini.
Définition 1.2 : *(hors programme)* ensemble infini.
Définition 1.3 : ensemble dénombrable.
Théorème 1.1 : énumération des éléments d'un ensemble fini ou dénombrable.
Théorème 1.2 : *(hors programme)* parties de $\mathbb{N}$.
Théorème 1.3 : *(hors programme)* caractérisation des ensembles finis ou dénombrables.
Théorème 1.4 : produit cartésien d'ensembles dénombrables.
Théorème 1.5 : *(hors programme)* réunion dénombrable d'ensembles au plus dénombrables.

2. Espaces probabilisés.

Définition 2.1 : tribu.
Théorème 2.1 : propriétés élémentaires d'une tribu.
Définition 2.2 : probabilité sur $(\Omega,\mathscr{A})$, espace probabilisé.
Théorème 2.2 : conséquences de la définition d'une probabilité.
Théorème 2.3 : probabilité d'une partie finie.
rappel : correspondance de vocabulaire.
Théorème 2.4 : continuité croissante et décroissante d'une probabilité.
Théorème 2.5 : probabilité d'une réunion d'évènements.

3. Probabilités conditionnelles.

Théorème 3.1 et Définition 3.1 : probabilité conditionnelle.
Théorème 3.2 : formule des probabilités composées.
Théorème 3.3 : généralisation de la formule des probabilités composées.
Définition 3.2 : indépendance d'évènements et indépendance mutuelle.
Théorème 3.4 : caractérisation de l'indépendance de deux évènements.
Théorème 3.5 : liens entre les notions d'indépendance.
Définition 3.3 : système complet dénombrable d'évènements.
Théorème 3.6 : formule des probabilités totales.
Définition 3.4 : événement presque sûr, événement négligeable.
Théorème 3.7 : généralisation (système quasi complet d'événements).
Théorème 3.8 : formule de Bayes.

Espaces probabilisés.

1. Ensembles dénombrables.

Définition 1.1 : ensemble fini

Soit E un ensemble.

On dit que E est fini si et seulement si il existe un entier : $n \in \mathbb{N}$, tel que E soit en bijection avec : $\mathbb{N}_n = \{1, \dots, n\}$.

Si : $n = 0$, $\mathbb{N}_n = \varnothing$, et cet ensemble caractérise l'ensemble vide.

Remarques :

- On admet que l'entier n précédent est unique (on montre que s'il existe une injection de N_n dans N_p, alors : $n \leq p$, et donc s'il existe une bijection de N_n dans $\mathbb{N}_p$ (par l'intermédiaire d'un ensemble fini E), alors : $n = p$).
- n est évidemment le cardinal de E.
- Une définition équivalente est de dire qu'il existe : $n \in \mathbb{N}$, tel que E soit en bijection avec $\{0, \dots, n - 1\}$.

En effet, l'application f de $\mathbb{N}_n$ dans $\{0, \dots, n - 1\}$, qui à k fait correspondre $k - 1$, est bijective.

Définition 1.2 : (hors programme) ensemble infini

Soit E un ensemble.

On dit que E est infini si et seulement si E n'est pas fini.

Remarque :

Une autre définition du fait que E est infini est de dire qu'il existe une bijection entre E et une de ses parties strictes (ou propres) E', c'est-à-dire telle que : $E' \subset E$, $E' \neq E$.

Exemples 1.2 :

- toute partie non majorée de $\mathbb{N}$ est infinie.
- $\mathbb{R}$ est infini car $\mathbb{R}$ est en bijection avec $\mathbb{R}^{+*}$ par l'exponentielle, et : $\mathbb{R}^{+*} \subset \mathbb{R}$, avec : $\mathbb{R}^{+*} \neq \mathbb{R}$.

Définition 1.3 : ensemble dénombrable

Soit E un ensemble.

On dit que E est dénombrable s'il existe une bijection de E dans $\mathbb{N}$.

Théorème 1.1 : énumération des éléments d'un ensemble fini ou dénombrable

Soit E un ensemble.

- Si E est fini, alors : $\exists\, n \in \mathbb{N}$, $E = \{x_1, \dots, x_n\}$, où les x_i sont distincts deux à deux.
- Si E est infini, alors : $E = \{x_k, k \in \mathbb{N}\}$, où les x_k sont distincts deux à deux.

démonstration :

- Si E est fini, soit φ une bijection de E dans $\mathbb{N}_n$, alors en notant : $\forall\, 1 \leq k \leq n$, $x_k = \varphi^{-1}(k)$, on constate que les x_k sont distincts deux à deux (puisque φ et φ^{-1} sont injectives) et que : $E = \{x_1, \dots, x_n\} = \{\varphi^{-1}(1), \dots, \varphi^{-1}(n)\}$, (puisque φ et φ^{-1} sont surjectives).
- Si E est dénombrable, on considère de même une bijection φ de E dans $\mathbb{N}$, et on note alors : $\forall\, n \in \mathbb{N}$, $x_n = \varphi^{-1}(n)$.

Comme dans le premier point, les x_n sont distincts deux à deux et : $E = \{\varphi^{-1}(n), n \in \mathbb{N}\} = \{x_n, n \in \mathbb{N}\}$, (puisque φ et φ^{-1} sont surjectives).

Remarques :

- Toute partie d'un ensemble fini est finie.
- Si E et F sont finis, alors $(E \cup F)$, $(E \cap F)$, $(E \setminus F)$ et $E \times F$ sont finis.

Théorème 1.2 : (hors programme) parties de $\mathbb{N}$

Soit A une partie de $\mathbb{N}$.

Alors A est finie ou dénombrable (on dit alors que A est « au plus dénombrable »).

démonstration :

Soit donc A une partie de $\mathbb{N}$, qu'on va supposer infinie.

On définit par récurrence l'application φ de la façon suivante :

• $\varphi(0) = \min(A)$, qui existe puisque A est une partie non vide de $\mathbb{N}$.

• $\forall n \geq 1$, on pose : $\varphi(n) = \min(A \setminus \{\varphi(0), \ldots, \varphi(n-1)\})$, autrement dit le plus petit élément de A une fois éliminés les n premiers plus petits.

Remarquons que $\varphi(n)$ existe pour tout n sinon A serait fini.

L'application φ est une bijection de $\mathbb{N}$ dans A.

En effet :

• elle est injective, car si : $p \neq q$ (par exemple : $p > q$), alors : $\varphi(p) \notin \{\varphi(0), \ldots, \varphi(q)\}$, et donc : $\varphi(p) \neq \varphi(q)$.

• elle est surjective.

Pour cela soit : $a \in A$.

Supposons que : $\forall n \in \mathbb{N}$, $\varphi(n) \neq a$.

On aurait alors : $\forall n \in \mathbb{N}^*$, $a \in A \setminus \{\varphi(0), \ldots, \varphi(n-1)\}$, et par définition de $\varphi(n)$: $\varphi(n) \leq a$.

On a de plus par définition de $\varphi(0)$: $\varphi(0) \leq a$.

Donc a serait un entier majorant une partie infinie de $\mathbb{N}$ ce qui est impossible.

Donc φ est injective et surjective et elle définit une bijection de $\mathbb{N}$ dans E.

E est donc alors dénombrable.

Remarque :

De même, toute partie d'un ensemble dénombrable est finie ou dénombrable (au plus dénombrable).

Théorème 1.3 : (hors programme) caractérisation des ensembles finis ou dénombrables

Soit E un ensemble.

E est fini ou dénombrable si et seulement si il existe une surjection de $\mathbb{N}$ dans E.

démonstration :

• La condition est nécessaire.

En effet, si E est dénombrable alors il existe une bijection de E dans $\mathbb{N}$ et donc une bijection de $\mathbb{N}$ dans E, (par exemple la réciproque de la précédente) qu'on peut noter φ.

φ est donc une application surjective de $\mathbb{N}$ dans E.

Si E est fini, il existe une bijection de E dans $\{0,\ldots, n-1\}$ donc en notant f sa réciproque, on peut alors construire φ de $\mathbb{N}$ dans E par :

$\forall 0 \leq k \leq n-1$, $\varphi(k) = f(k)$, puis : $\forall k \geq n$, $\varphi(k) = f(0)$.

L'application φ ainsi construite est clairement surjective de $\mathbb{N}$ dans E.

• La condition est suffisante.

En effet, si f est une surjection de $\mathbb{N}$ dans E, alors :

$\forall x \in E$, $f^{-1}(x)$ est une partie non vide de E (puisque f est surjective).

On pose alors : $\forall x \in E$, $m(x) = \min\{n \in \mathbb{N}, f(n) = x\}$.

On dispose ainsi d'une application m de E dans $\mathbb{N}$, et par construction : $f o m = id_E$.

Donc m est injective.

m est donc bijective de E dans : $m(E) \subset \mathbb{N}$.

E est donc en bijection avec une partie de $\mathbb{N}$ et E est donc finie ou dénombrable.

Théorème 1.4 : produit cartésien d'ensembles dénombrables

Soient E et F deux ensembles dénombrables.

Alors $E\times F$, ensemble des couples formés d'un élément de E et d'un élément de F, est dénombrable.

démonstration :

• Commençons avec : $E = F = \mathbb{N}^2$, et montrons qu'il existe une bijection de $\mathbb{N}^2$ dans $\mathbb{N}$.

On utilise pour cela ce que l'on appelle l'énumération diagonale.

On pose pour cela : $\forall (a,b) \in \mathbb{N}^2$, $f(a,b) = b + (0 + 1 + \ldots + (a+b))$.

On pourra essayer de représenter $\mathbb{N}^2$ sur des axes Ox et Oy d'un repère orthonormé et indiquer en chaque point (a,b) l'image f(a,b) pour comprendre l'origine de « énumération diagonale ».

f ainsi construite est bijective.

Pour cela : $\forall n \in \mathbb{N}$, on note : $u_n = 0 + 1 + \ldots + n$, et on remarque que (u_n) est strictement croissante.

Soit maintenant : $p \in \mathbb{N}$.

- S'il existe un couple : (a,b) ∈ $\mathbb{N}^2$, tel que : $p = f(a,b) = (0+1+\ldots+(a+b)) + b = u_{a+b} + b$,
alors avec : $m = a+b$, on a : $u_m \le u_m + b = p \le u_m + (a+b) = u_m + m < u_m + (m+1) = u_{m+1}$.
D'autre part : $b = p - u_m$, et : $a = m - b$.
- Réciproquement, (u_n) étant strictement croissante, on sait que : ∃ ! m ∈ $\mathbb{N}$, tel que : $u_m \le p < u_{m+1}$.
Posons alors : $b = p - u_m$, et : $a = m - b$ (unique possibilité).
On constate ensuite que : b ≥ 0 (par définition de m), puis : $b = p - u_m < u_{m+1} - u_m = m+1$, autrement dit : $b \le m$, et donc : a ≥ 0.
Autrement dit, l'unique couple (a,b) ainsi trouvé est un élément de $\mathbb{N}^2$ et par construction, on a bien :
$f(a,b) = b + u_{a+b} = b + u_m = p$.
f étant bijective de $\mathbb{N}^2$ dans $\mathbb{N}$, $\mathbb{N}^2$ est donc dénombrable.
• Considérons maintenant E et F deux ensembles dénombrables.
Soit f une bijection de E dans $\mathbb{N}$, et g une bijection de F dans $\mathbb{N}$.
Alors l'application h définie par : ∀ (x,y) ∈ E×F, h((x,y)) = (f(x),g(y)), est une bijection de E×F dans $\mathbb{N}^2$, et par composition, on en déduit qu'il existe une bijection de E×F dans $\mathbb{N}$.
E×F est donc dénombrable.

Théorème 1.5 : (hors programme) réunion dénombrable d'ensembles au plus dénombrables
Soit $(E_n)_{n\in\mathbb{N}}$ une famille dénombrable d'ensembles finis ou dénombrables.
Alors la réunion : $E = \bigcup_{n=0}^{+\infty} E_n$, est finie ou dénombrable.

Démonstration :
Chaque ensemble E_n étant fini ou dénombrable, il existe une bijection f_n de $\mathbb{N}$ dans E_n.
Considérons alors l'application φ de $\mathbb{N}\times\mathbb{N}$ dans E définie par :
∀ (n,p) ∈ $\mathbb{N}\times\mathbb{N}$, φ(n,p) = f_n(p).
φ est surjective.
En effet : ∀ a ∈ E, ∃ n ∈ $\mathbb{N}$, a ∈ E_n, et : ∃ p ∈ $\mathbb{N}$, a = f_n(p), autrement dit : a = φ(n,p).
Notons enfin ψ une bijection de $\mathbb{N}$ dans $\mathbb{N}\times\mathbb{N}$.
L'application φoψ est alors surjective de $\mathbb{N}$ dans E et E est fini ou dénombrable.

2. Espaces probabilisés.

Définition 2.1 : tribu
Soit Ω un ensemble.
Une famille $\mathcal{A}$ de parties de Ω (soit un sous-ensemble de $\mathcal{P}(\Omega)$) est appelée tribu sur Ω si et seulement si elle vérifie les propriétés suivantes :
- Ω ∈ $\mathcal{A}$,
- ∀ A ∈ $\mathcal{A}$, $\overline{A}$ ∈ $\mathcal{A}$,
- ∀ $(A_n)_{n\in\mathbb{N}}$ ∈ $\mathcal{A}^{\mathbb{N}}$, $\bigcup_{n\notin N} A_n \in \mathcal{A}$.

Théorème 2.1 : propriétés élémentaires d'une tribu
Si $\mathcal{A}$ est une tribu sur Ω, alors :
- ∅ ∈ $\mathcal{A}$,
- ∀ $(A_i)_{1\le i\le n}$ ∈ $\mathcal{A}^n$, $\bigcup_{1\le i\le n} A_n \in \mathcal{A}$,
- ∀ $(A_i)_{1\le i\le n}$ ∈ $\mathcal{A}^n$, $\bigcap_{1\le i\le n} A_n \in \mathcal{A}$,
- ∀ $(A_n)_{n\in\mathbb{N}}$ ∈ $\mathcal{A}^{\mathbb{N}}$, $\bigcap_{n\notin N} A_n \in \mathcal{A}$,
- ∀ (A,B) ∈ $\mathcal{A}^2$, A ∩ B ∈ $\mathcal{A}$.

démonstration :

- Ω est dans $\mathcal{A}$ et $\mathcal{A}$ est stable par passage au complémentaire et : $\emptyset = \overline{\Omega}$, donc : $\emptyset \in \mathcal{A}$.
- Si $(A_i)_{1\leq i\leq n}$ est une famille finie d'éléments de $\mathcal{A}$, il suffit de poser : $\forall\ 1 \leq i \leq n$, $B_i = A_i$, et : $\forall\ i \geq n + 1$, $B_i = \Omega$, pour avoir : $\bigcup\limits_{1\leq i\leq n} A_n = \bigcup\limits_{i\in N} B_i$, et comme tous les B_i sont dans $\mathcal{A}$, on a donc : $\bigcup\limits_{1\leq i\leq n} A_n \in \mathcal{A}$.
- Il suffit de remarquer que : $\forall\ (A_i)_{1\leq i\leq n} \in \mathcal{A}^n$, $\bigcap\limits_{1\leq i\leq n} A_n = \overline{\bigcup\limits_{1\leq i\leq n} \overline{A_n}}$, et : $\bigcap\limits_{1\leq i\leq n} A_n \in \mathcal{A}$.
- De même : $\forall\ (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$, $\bigcap\limits_{n\in N} A_n = \overline{\bigcup\limits_{n\in N} \overline{A_n}}$, et : $\bigcap\limits_{n\in N} A_n \in \mathcal{A}$.
- Ce dernier point est un cas particulier du point 3 (pour deux éléments de $\mathcal{A}$).

Remarques :

- La notion de tribu permet d'étendre le cadre $\mathcal{P}(\Omega)$ utilisé pour définir des probabilités sur les ensembles finis.
- L'ensemble Ω est toujours censé servir de cadre à la modélisation d'une expérience aléatoire et $\mathcal{A}$ représente l'ensemble des événements envisagés pour le résultat de cette expérience.

Définition 2.2 : probabilité sur $(\Omega,\mathcal{A})$, espace probabilisé

Soit Ω un ensemble et $\mathcal{A}$ une tribu sur Ω.

On appelle probabilité sur $(\Omega,\mathcal{A})$ une application P de $\mathcal{A}$ dans [0,1] telle que :

- $P(\Omega) = 1$,
- pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ disjoints deux à deux (événements deux à deux incompatibles), la série $\sum\limits_{n\geq 0} P(A_n)$ est convergente et : $P(\bigcup\limits_{n\in N} A_n) = \sum\limits_{n=0}^{+\infty} P(A_n)$.

Cette dernière propriété est appelée σ-additivité de P.

Un tel triplet $(\Omega,\mathcal{A},P)$ est appelé espace probabilisé et les éléments de $\mathcal{A}$ sont appelés événements.

Théorème 2.2 : conséquences de la définition d'une probabilité

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

- $P(\emptyset) = 0$,
- Soit : $(A_i)_{1\leq i\leq n} \in \mathcal{A}^n$, tel que les éléments (A_i) sont deux à deux disjoints.

  Alors : $P\left(\bigcup\limits_{i=1}^{n} A_i\right) = \sum\limits_{i=1}^{n} P(A_i)$.
- $\forall\ A \in \mathcal{A}$, $P(\overline{A}) = 1 - P(A)$.
- $\forall\ (A,B) \in \mathcal{A}^2$, $(A \subset B) \Rightarrow (P(A) \leq P(B))$.
- $\forall\ (A,B) \in \mathcal{A}^2$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- en particulier : $\forall\ (A,B) \in \mathcal{A}^2$, $P(A \cup B) \leq P(A) + P(B)$.
- $\forall\ (A,B) \in \mathcal{A}^2$, $(A \cap B = \emptyset) \Rightarrow (P(A \cup B) = P(A) + P(B))$.
- $\forall\ (A,B) \in \mathcal{A}^2$, $(B \subset A) \Rightarrow (P(A \setminus B) = P(A) - P(B))$.

démonstration :

- On pose : $\forall\ n \in \mathbb{N}$, $A_n = \emptyset$, et la famille (A_n) est une famille d'éléments de $\mathcal{A}$ deux à deux disjoints. Donc la série $\sum\limits_{n\geq 0} P(A_n)$ converge, mais tous les termes étant égaux, cette série est la série nulle d'où on déduit que : $\forall\ n \in \mathbb{N}$, $P(A_n) = P(\emptyset) = 0$.
- Il suffit de construire : $\forall\ 1 \leq i \leq n$, $B_i = A_i$, et : $\forall\ n + 1 \leq i$, $B_i = \emptyset$.

  La famille (B_i) est une famille d'éléments de $\mathcal{A}$ deux à deux disjoints donc la série $\sum\limits_{n\geq 1} P(B_n)$ converge et : $P\left(\bigcup\limits_{i=1}^{n} A_i\right) = P(\bigcup\limits_{n\in N^*} B_n) = \sum\limits_{n=1}^{+\infty} P(B_n) = \sum\limits_{i=1}^{n} P(A_i)$.

- Puisque : $A \cap \bar{A} = \emptyset$, on a : $P(A \cup \bar{A}) = P(\Omega) = P(A) + P(\bar{A})$, soit : $P(\bar{A}) = 1 - P(A)$.
- Pour : $A \subset B$, on peut écrire : $B = A \cup (B \setminus A)$, avec : $A \cap (B \setminus A) = \emptyset$.
  Donc : $P(B) = P(A) + P(B \setminus A)$, et P étant à valeurs positives, on déduit : $P(B) \geq P(A)$.
- On part ensuite de :
  $A = (A \setminus B) \cup (A \cap B)$, union disjointe, et : $P(A) = P(A \setminus B) + P(A \cap B)$,
  $B = (B \setminus A) \cup (A \cap B)$, union disjointe, et : $P(B) = P(B \setminus A) + P(A \cap B)$,
  $(A \cup B) = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$, et : $P(A \cup B) = P(A \setminus B) + P(A \cap B) + P(B \setminus A)$,
  et en soustrayant les deux premières égalités à la troisième, on obtient :
  $P(A \cup B) - P(A) - P(B) = - P(A \cap B)$, d'où le résultat.
- Puisque $P(A \cap B)$ est positive, on en déduit le dernière résultat.
- On retrouve avec : $A \cap B = \emptyset$, le deuxième point dans le cas de deux ensembles disjoints.
- il suffit de remarquer que : $(B \subset A) \Rightarrow (A \cap (A \setminus B) = \emptyset$, et : $A = B \cup (A \setminus B))$, et donc :
  $P(B) + P(A \setminus B) = P(A)$.

remarque :

Ces résultats sont identiques à ceux obtenus dans le cadre des espaces probabilisés finis.

Théorème 2.3 : probabilité d'une partie finie

Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé.

Si : $A \in \mathscr{A}$, et A est une partie finie de Ω, alors : $P(A) = \sum_{\omega \in A} P(\omega)$.

démonstration :

Si A est fini, notons : $A = \{\omega_1, \ldots, \omega_n\}$.

Il suffit alors de considérer la famille finie d'ensembles disjoints : $\forall\ 1 \leq k \leq n$, $A_k = \{\omega_k\}$, et :

$$P(A) = P\left(\bigcup_{k=1}^{n} \{\omega_k\}\right) = \sum_{k=1}^{n} P(\{\omega_k\}) = \sum_{\omega \in A} P(\omega).$$

Rappel : correspondance de vocabulaire :

probabiliste $\Leftrightarrow$ ensembliste

résultat possible de l'expérience	ω, élément de Ω
évènement	$A \in \mathscr{A}$
A est réalisé	$\omega \in A$
A implique B	$A \subset B$
A ou B	$A \cup B$
A et B	$A \cap B$
contraire de A (A ne se produit pas)	$\bar{A}$ (complémentaire de A dans Ω)
évènement impossible	$\emptyset$
évènement certain	Ω
A et B sont incompatibles	$A \cap B = \emptyset$

Théorème 2.4 : continuité croissante et décroissante d'une probabilité

Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé et soit : $(A_n) \in \mathscr{A}^{\mathbb{N}}$.

- Si : $\forall\ n \in \mathbb{N}$, $A_n \subset A_{n+1}$, alors la suite $(P(A_n))$ converge et : $\lim_{n \to +\infty} P(A_n) = P(\bigcup_{n \in N} A_n)$.
- Si : $\forall\ n \in \mathbb{N}$, $A_{n+1} \subset A_n$, alors la suite $(P(A_n))$ converge et : $\lim_{n \to +\infty} P(A_n) = P(\bigcap_{n \in N} A_n)$.

démonstration :

- Pour le premier point, on remarque tout d'abord que : $\forall\ n \in \mathbb{N}$, $A_n \subset A_{n+1}$, on a : $P(A_n) \leq P(A_{n+1})$, et donc la suite $(P(A_n))$ est croissante.
  Etant de plus majorée par 1, elle converge.
  Notons maintenant : $B_0 = A_0$, et : $\forall\ n \in \mathbb{N}^*$, $B_n = A_n \setminus A_{n-1}$.
  Alors la famille (B_n) est une famille d'ensembles deux à deux disjoints.

De plus : ∀ n ∈ ℕ, $\bigcup_{i=0}^{n} B_i = A_n = \bigcup_{i=0}^{n} A_i$, et : ∀ n ∈ ℕ, $P(A_n) = P\left(\bigcup_{i=0}^{n} B_i\right) = \sum_{i=0}^{n} P(B_i)$.

Puis par double inclusion, on a : $\bigcup_{i=0}^{+\infty} B_i = \bigcup_{i=0}^{+\infty} A_i$.

En effet :

- ∀ ω ∈ $\bigcup_{i=0}^{+\infty} B_i$, ∃ j ∈ ℕ, ω ∈ $B_j \subset A_j$, et : ω ∈ $A_j \subset \bigcup_{i=0}^{+\infty} A_i$.
- ∀ ω ∈ $\bigcup_{i=0}^{+\infty} A_i$, soit : n = min{i ∈ ℕ, ω ∈ A_i}.

Si : n = 0, ω ∈ $A_0 = B_0$, et : ω ∈ $\bigcup_{i=0}^{+\infty} B_i$,

Si : n > 0, ω ∈ A_n, et : ω ∉ A_{n-1}, donc : ω ∈ B_n, et : ω ∈ $\bigcup_{i=0}^{+\infty} B_i$.

Enfin la définition d'une probabilité permet d'écrire :

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

- Pour le deuxième point, on peut raisonner par complémentaires.

Plus précisément on note : ∀ n ∈ ℕ, $A'_n = \overline{A_n}$, la suite (A'_n) est alors une suite croissante d'éléments de $\mathscr{A}$ (au sens de l'inclusion) et : $\bigcap_{n\in N} A_n = \overline{\bigcup_{n\in N} A'_n}$.

On utilise alors le point 1 et les propriétés d'une probabilité pour écrire :

- la suite ($P(A'_n)$) converge donc la suite ($P(A_n)$) aussi puisque : ∀ n ∈ ℕ, $P(A_n) = 1 - P(A'_n)$,
- $1 - \lim_{n\to+\infty} P(A_n) = \lim_{n\to+\infty} [1 - P(A_n)] = \lim_{n\to+\infty} P(A'_n) = P(\bigcup_{n\in N} A'_n) = 1 - P\left(\bigcap_{n\in N} A_n\right)$,

d'où le résultat voulu.

Théorème 2.5 : probabilité d'une réunion d'évènements

Soit (Ω,$\mathscr{A}$,P) un espace probabilisé.

- Si $(A_n)_{1\le i\le n}$ est une famille finie d'éléments de $\mathscr{A}$, alors :
  $P(A_1 \cup \ldots \cup A_n) \le P(A_1) + \ldots + P(A_n)$.
- Si de plus : $(A_n) \in \mathscr{A}^{\mathbb{N}}$, et si la série $\sum_{n\ge 0} P(A_n)$ converge, alors : $P(\bigcup_{n\in N} A_n) \le \sum_{n=0}^{+\infty} P(A_n)$.

démonstration :

- On démontre évidemment le premier résultat par récurrence sur n.

Pour une seule partie A_1, c'est immédiat.

Si on suppose le résultat établi pour n parties de Ω (avec : n ≥ 1), on peut alors considérer (n+1) parties de Ω et écrire : $P(A_1 \cup \ldots \cup A_{n+1}) \le P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) \le P(A_1) + \ldots + P(A_n) + P(A_{n+1})$, en utilisant le théorème 2.2 (pour la première inégalité).

- Pour le deuxième résultat, on peut poser : ∀ n ∈ ℕ, $B_n = \bigcup_{k=0}^{n} A_k$, et on a : $\bigcup_{k=0}^{+\infty} B_k = \bigcup_{k=0}^{+\infty} A_k$.

La suite (B_n) est croissante donc la suite ($P(B_n)$) converge et : $\lim_{n\to+\infty} P(B_n) = P\left(\bigcup_{n\in N} B_n\right) = P(\bigcup_{n\in N} A_n)$.

De plus : ∀ n ∈ ℕ, $P(B_n) = P\left(\bigcup_{k=0}^{n} A_k\right) \le P(A_0) + \ldots + P(A_n)$.

Si maintenant, on fait tendre n vers +∞, on conclut que : $P(\bigcup_{n\in N} A_n) = \lim_{n\to+\infty} P(B_n) \le \sum_{n=0}^{+\infty} P(A_n)$.

3. Probabilités conditionnelles.

Théorème 3.1 et définition 3.1 : probabilité conditionnelle

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

Si B est un évènement tel que : $P(B) > 0$, alors l'application P_B définie par :

$\forall A \in \mathcal{A}$, $P_B(A) = \dfrac{P(A\cap B)}{P(B)}$,

est une probabilité sur $(\Omega,\mathcal{A})$, appelée probabilité conditionnelle sachant B.

On notera également : $P_B(A) = P(A|B)$, et on lira « probabilité de A sachant B ».

démonstration :

• P_B est correctement définie sur $\mathcal{A}$ et à valeurs dans [0,1] car :

$\forall A \in \mathcal{A}$, $P_B(A) = \dfrac{P(A\cap B)}{P(B)} \geq 0$, et : $P_B(A) = \dfrac{P(A\cap B)}{P(B)} \leq \dfrac{P(B)}{P(B)} = 1$, puisque : $A \cap B \subset B$.

• $P_B(\Omega) = \dfrac{P(\Omega\cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$.

• $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$, famille d'événements deux à deux disjoints, notons :

$\forall n \in \mathbb{N}$, $A'_n = A_n \cap B$.

Alors la famille (A'_n) est une famille d'événements (de $\mathcal{A}$) deux à deux disjoints, car :

$\forall (n,p) \in \mathbb{N}^2$, $(n \neq p) \Rightarrow ((A'_n \cap A'_p = (A_n \cap B) \cap (A_p \cap B) \subset (A_n \cap A_p) = \emptyset)$.

Donc la série : $\sum_{n\geq 0} P(A'_n) = \sum_{n\geq 0} P(A_n\cap B)$, converge tout comme $\sum_{n\geq 0}\dfrac{P(A_n\cap B)}{P(B)}$, et :

$$P((\bigcup_{n\notin N} A_n)\cap B) = P(\bigcup_{n\notin N}(A_n\cap B)) = P(\bigcup_{n\notin N} A'_n) = \sum_{n=0}^{+\infty} P(A'_n) = \sum_{n=0}^{+\infty} P(A_n\cap B).$$

On en déduit que : $P_B(\bigcup_{n\notin N} A_n) = \dfrac{P((\bigcup_{n\notin N} A_n)\cap B)}{P(B)} = \sum_{n=0}^{+\infty}\dfrac{P(A_n\cap B)}{P(B)} = \sum_{n=0}^{+\infty} P_B(A_n)$.

Convention :

Si $(\Omega,\mathcal{A},P)$ est un espace probabilisé et si : $B \in \mathcal{A}$, avec : $P(B) = 0$, alors on convient de noter :

$\forall A \in \mathcal{A}$, $P_B(A) = P(A|B) = 0$.

Théorème 3.2 : formule des probabilités composées

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

Soient A et B deux évènements de $\mathcal{A}$.

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

démonstration :

Supposons dans un premier temps que : $P(A) > 0$, et : $P(B) > 0$.

Il suffit alors de reprendre la définition des probabilités conditionnelles en échangeant les rôles de A et de B, ce qui donne :

$P_B(A) = \dfrac{P(A\cap B)}{P(B)}$, donc :

$P(A\cap B) = P_B(A).P(B)$, et de façon symétrique : $P(A\cap B) = P_A(B).P(A)$.

Si maintenant : $P(A) = 0$, et : $P(B) > 0$, alors : $P_A(B) = P(B|A) = 0$ (par convention) et :

$P(A \cap B) = 0$, car : $A \cap B \subset A$, donc : $P_B(A) = \dfrac{P(A\cap B)}{P(B)} = 0$, d'où à nouveau l'égalité.

Enfin si P(A) et P(B) sont nulles, la convention précédente vaut pour les deux termes et conduit une fois de plus à l'égalité voulue.

Théorème 3.3 : généralisation de la formule des probabilités composées

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

Soit : $n \geq 2$, et $A_1, \ldots, A_n$ des événements de $\mathscr{A}$ tels que : $P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq 0$, alors :

• $\forall\ 1 \leq j \leq n-1$, $P\left(\bigcap_{i=1}^{j} A_i\right) \neq 0$, et :

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

démonstration :

On commence par remarquer que : $\forall\ 1 \leq j \leq n-1$, $\bigcap_{i=1}^{n-1} A_i \subset \bigcap_{i=1}^{j} A_i$, on a : $P\left(\bigcap_{i=1}^{j} A_i\right) \neq 0$.

Puis on obtient l'égalité voulue par récurrence à partir du théorème 3.2.

Définition 3.2 : indépendance d'évènements et indépendance mutuelle

Soit $(\Omega,\mathscr{A},P)$, un espace probabilisé.

Deux évènements A et B de $\mathscr{A}$ sont dits indépendants si et seulement si :

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

Les évènements d'une famille $(A_j)_{j\in J}$ quelconque de $\mathscr{A}$ sont dits indépendants dans leur ensemble ou mutuellement indépendants si et seulement si, pour toute partie finie : $J \subset I$, on a :

$P(\bigcap_{i\in J} A_i) = \prod_{i\in J} P(A_i)$.

En particulier, si $(A_i)_{1\leq i\leq n}$ est une famille finie d'évènements, ils sont indépendants si et seulement si :

$\forall\ 1 \leq j_1 < \ldots < j_k \leq n$, on a : $P(A_{j_1} \cap ... \cap A_{j_k}) = P(A_{j_1})...P(A_{j_k})$.

Théorème 3.4 : caractérisation de l'indépendance de deux évènements

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé et soit : $(A,B) \in \mathscr{A}^2$.

Si : $P(B) > 0$, A et B sont indépendants si et seulement si : $P(A|B) = P(A)$.

démonstration :

Ce résultat est immédiat car si : $P(B) \neq 0$, alors : $P_B(A) = \dfrac{P(A \cap B)}{P(B)}$.

Donc A et B sont indépendants si et seulement si : $P(A \cap B) = P(A).P(B)$, soit le résultat.

Remarque :

La lecture « naïve » de cette propriété traduit l'idée « intuitive » qu'on se fait de l'indépendance : si A et B sont indépendants, la probabilité de voir A se produire est la même que celle de voir A se produire connaissant la réalisation ou pas de B, autrement dit B n'influe pas sur la probabilité de voir A se produire.

Théorème 3.5 : liens entre les notions d'indépendance

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé.

• si : $n = 2$, dans la définition précédente, on retrouve la première définition.

• la condition d'indépendance pour une famille finie d'évènements correspond à $2^n - n - 1$ égalités.

• si des évènements sont mutuellement indépendants, ils sont indépendants deux à deux, mais la réciproque est fausse.

démonstration :

• si l'ensemble d'indices J considéré comporte 0 indice (un ensemble : $J = \emptyset$) ou 1 indice (un singleton), la condition d'indépendance est toujours satisfaite (soit en tout 0 égalités).

Donc l'indépendance de n événements correspond bien à 2^n (nombre de familles d'indices possibles) moins $(n + 1)$ familles d'indices et donc $(2^n - n - 1)$ conditions à satisfaire.

• pour : $n = 2$, et donc : $A_1 = A$, $A_2 = B$, il y a une seule condition à satisfaire correspondant à :

$P(A_1 \cap A_2) = P(A \cap B) = P(A).P(B) = P(A_1).P(A_2)$.

• supposons que $(A_i)_{1\leq i\leq n}$ soit une famille de n événements mutuellement indépendants.

Alors pour tout couple (A_i, A_j), avec : $i \neq j$, on utilise l'ensemble d'indices : $J = \{i,j\}$, et on réécrit l'égalité

pour cet ensemble J qui donne : $P(A_i \cap A_j) = P(A_i).P(A_j)$, soit bien l'égalité qui montre l'indépendance de A_i et de A_j.
- si on considère un ensemble : $\Omega = \{a,b,c,d\}$, muni de la probabilité uniforme et les événements : $A = \{a,d\}$, $B = \{b,d\}$, et : $C = \{c,d\}$,

alors : $P(A) = P(B) = P(C) = \frac{1}{2}$.

Puis : $P(A \cap B) = P(\{d\}) = \frac{1}{4} = P(A).P(B)$, de même pour $P(A \cap C)$ et $P(B \cap C)$.

Remarque :

On montre (voir exercices) que si A et B sont indépendants alors $\overline{A}$ et B le sont aussi, tout comme $\overline{A}$ et $\overline{B}$.

Définition 3.3 : système complet dénombrable d'évènements

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé.

On dit que $(A_n)_{n\in\mathbb{N}}$ est un système complet dénombrable d'évènements si et seulement si :
- $\forall (i,j) \in \mathbb{N}^2$, $(i \neq j) \Rightarrow (A_i \cap A_j = \emptyset)$, (incompatibles deux à deux),
- $\bigcup_{n\in N} A_n = \Omega$.

Cela correspond à une partition de l'ensemble Ω.

Remarques :
- Cela vient en complément des systèmes complets finis d'événements.
- Dans le cas d'un système complet dénombrable d'événements, on peut noter :

$\forall n \in \mathbb{N}$, $B_n = \bigcup_{i=0}^{n} A_i$, et la suite (B_n) est croissante.

Donc : $\lim_{n\to+\infty} P(B_n) = P(\bigcup_{n\in N} A_n) = P(\Omega) = 1 = \lim_{n\to+\infty} \sum_{i=0}^{n} P(A_i)$,

puisque les A_i sont disjoints deux à deux.

Autrement dit la série $\sum_{n\geq 0} P(A_n)$ converge et sa somme vaut 1.

Théorème 3.6 : formule des probabilités totales

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ un système complet d'évènements telle que : $\forall n \in \mathbb{N}$, $P(A_n) > 0$.

Alors pour tout : $B \in \mathscr{A}$, la série $\sum_{n\geq 0} P(B \cap A_n)$ converge et :

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

démonstration :

On commence par remarquer que :

$\forall (n,p) \in \mathbb{N}^2$, $(n \neq p) \Rightarrow ((B \cap A_n) \cap (B \cap A_p) \subset (A_n \cap A_p) = \emptyset)$.

Donc les événements $(B \cap A_n)$ sont deux à deux disjoints et :

$$P\left(\bigcup_{n\geq 0}(B \cap A_n)\right) = P\left(B \cap (\bigcup_{n\geq 0} A_n)\right) = P(B \cap \Omega) = P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P(B|A_n).P(A_n).$$

La dernière égalité correspond simplement à une réécriture de la quantité précédente.

Définition 3.4 : événement presque sûr, événement négligeable

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé.

On dit qu'un événement A de $\mathscr{A}$ est presque sûr lorsque : $P(A) = 1$.

On dit à l'inverse qu'un événement A de $\mathscr{A}$ est négligeable lorsque : $P(A) = 0$.

Le complémentaire d'un événement presque sûr est donc un événement négligeable.

Théorème 3.7 : généralisation (système quasi complet d'événements)

Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$, une famille d'évènements deux à deux incompatibles, telle que : $\sum_{n=0}^{+\infty} P(A_n) = 1$.

Alors le résultat précédent reste valable, autrement dit pour : $B \in \mathcal{A}$, la série $\sum_{n\geq 0} P(B\cap A_n)$ converge et : $P(B) = \sum_{n=0}^{+\infty} P(B|A_n).P(A_n) = \sum_{n=0}^{+\infty} P(B\cap A_n) = \sum_{n=0}^{+\infty} P_{A_n}(B).P(A_n)$.

démonstration :

On reprend la démonstration du théorème précédent, mais on note : $A = \bigcup_{n\geq 0} A_n$.

Alors : $P\left(\bigcup_{n\geq 0}(B\cap A_n)\right) = P\left(B\cap(\bigcup_{n\geq 0} A_n)\right) = P(B\cap A) = \sum_{n=0}^{+\infty} P(B\cap A_n) = \sum_{n=0}^{+\infty} P(B|A_n).P(A_n)$.

On remarque ensuite que $(A,\overline{A})$ forme un système complet d'événements et :

$P(B) = P(B\cap A) + P(B\cap \overline{A})$.

Or : $B\cap\overline{A} \subset \overline{A}$, et : $P(\overline{A}) = 1 - P(A) = 1 - 1 = 0$, donc : $P(B\cap\overline{A}) = 0$, et : $P(B) = P(B\cap A)$.

Ceci permet de conclure à l'égalité voulue.

Remarque :

La situation précédente correspond au cas où $\bigcup_{n=0}^{+\infty} A_n$ est un événement certain, son complémentaire étant un événement négligeable.

Théorème 3.8 : formule de Bayes

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

- Si A et B sont des événements de probabilité non nulle, alors : $P(A|B) = \dfrac{P(A).P(B|A)}{P(B)}$.
- Si de plus $(A_n)_{n\in\mathbb{N}}$ un système complet (ou quasi complet) d'évènements tel que : $\forall n \in \mathbb{N}$, $P(A_n) > 0$, alors :

  $\forall B \in \mathcal{A}$, tel que : $P(B) > 0$, on a : $\forall n \in \mathbb{N}$, $P(A_n|B) = \dfrac{P(A_n).P(B|A_n)}{\sum_{i=0}^{+\infty} P(A_i).P(B|A_i)}$.

démonstration :

- La démonstration du premier point est immédiate, puisque par définition, pour : $P(B) \neq 0$, on a :

  $P_B(A) = P(A|B) = \dfrac{P(A\cap B)}{P(B)}$, et donc : $P(A \cap B) = P(B).P(A|B)$.

  De même pour : $P(A) \neq 0$, on a : $P(A \cap B) = P(A).P(B|A)$.

  D'où le résultat avec l'égalité qui en découle.

- Le deuxième point s'obtient en remarquant simplement que : $P(B) = \sum_{i=0}^{+\infty} P(B|A_i).P(A_i)$, par la formule des probabilités totales.