

Ensemble des événements et probabilité.

On cherche à modéliser une expérience aléatoire, on nomme Ω l'ensemble des résultats (l'univers).

1.1 Ensemble des événements.

1.1.1 Notion de tribu.

Soit \mathcal{T} une partie de $\mathcal{P}(\Omega)$,
dire que \mathcal{T} est une **tribu** sur Ω signifie que :

- ❶ $\Omega \in \mathcal{T}$ (Contient l'événement certain)
- ❷ $\forall A \in \mathcal{T}, \quad \bar{A} \in \mathcal{T}$. (\mathcal{T} est stable par passage au complémentaire).
- ❸ $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \quad \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$. (\mathcal{T} est stable par union dénombrable).

Remarques : • Cette définition ne fera pas l'objet de questions directes, mais elle doit être connue.

- La donnée de (Ω, \mathcal{T}) est appelée : espace probabilisable.
- \mathcal{T} est l'ensemble des événements.
- Un ensemble E est dénombrable s'il existe une bijection de \mathbb{N} dans E .

autrement dit : s'il existe une suite (x_n) telle que $\begin{cases} E = \{ x_n \mid n \in \mathbb{N} \} \\ \text{et } \forall (i, j) \in \mathbb{N}^2, i \neq j \implies x_i \neq x_j \end{cases}$

Avec cette définition les ensembles dénombrables sont infinis. Vous trouverez d'autres définitions.

Propositions.

Soit \mathcal{T} une partie de $\mathcal{P}(\Omega)$,
si \mathcal{T} est une **tribu** alors :

- ❶ $\emptyset \in \mathcal{T}$ (Contient l'événement impossible).
- ❷ $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \quad \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$. (\mathcal{T} est stable par intersection dénombrable).
- ❸ \mathcal{T} est stable par réunion finie et par intersection finie.

Démonstrations.

Propriétés des intersections et des réunions dénombrables.

Soient $(A_n)_{n \in \mathbb{N}}$ une suite d'événements, B un événement et N un entier naturel non nul,

$$\text{Complémentaire : } \overline{\bigcap_{n=0}^{+\infty} A_n} = \bigcup_{n=0}^{+\infty} \overline{A_n} \quad \overline{\bigcup_{n=0}^{+\infty} A_n} = \bigcap_{n=0}^{+\infty} \overline{A_n}$$

$$\text{Distributivité, } B \cup \left(\bigcap_{n=0}^{+\infty} A_n \right) = \bigcap_{n=0}^{+\infty} (B \cup A_n) \quad B \cap \left(\bigcup_{n=0}^{+\infty} A_n \right) = \bigcup_{n=0}^{+\infty} (B \cap A_n).$$

$$\text{Inclusion, } \bigcap_{n=0}^{+\infty} A_n \subset \bigcap_{n=0}^N A_n \quad \bigcup_{n=0}^N A_n \subset \bigcup_{n=0}^{+\infty} A_n$$

1.1.2 Système complet d'événements

Définition.

Soit $(A_n)_{n \in I}$ une suite finie ou dénombrable d'événements ($I = \underbrace{[1; n]}_{\text{finie}}$ ou $I = \underbrace{\mathbb{N}}_{\text{dénombrable}}$),

Dire que $(A_n)_{n \in I}$ est un système complet d'événements signifie que :

$$\textcircled{1} \quad \forall (i, j) \in I^2, \quad i \neq j \implies A_i \cap A_j = \emptyset \quad \text{et} \quad \textcircled{2} \quad \bigcup_{n \in I} A_n = \Omega$$

Remarque : **1** se dit : (A_n) est une suite d'événements deux à deux disjoints.

ou encore : (A_n) est une suite d'événements deux à deux incompatibles.

1.2 Probabilité.

Définition.

Soient (Ω, \mathcal{T}) un espace probabilisable et P une application de \mathcal{T} dans \mathbb{R} .

Dire que P est une **probabilité** sur (Ω, \mathcal{T}) signifie que :

$$\textcircled{1} \quad \forall A \in \mathcal{T}, \quad P(A) \in [0, 1].$$

$$\textcircled{2} \quad P(\Omega) = 1$$

3 Quel que soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements. (*Axiome de σ -additivité*)

si les A_n sont deux à deux incompatibles,

$$\text{alors la série } \sum_{n \geq 0} P(A_n) \text{ converge et : } P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$$

Remarque : **3** entraîne : Si $A \cap B = \emptyset$ alors $P(A \cup B) = P(A) + P(B)$.

On retrouve donc les mêmes résultats que dans le cours de première année lorsque Ω est fini et $\mathcal{T} = \mathcal{P}(\Omega)$:

- n étant fixé dans \mathbb{N} , quel que soit $(A_k)_{1 \leq k \leq n}$ une suite de n événements,
si les A_k sont deux à deux incompatibles, alors $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=0}^n P(A_k)$
- $\forall A \in \mathcal{T}, \quad P(\overline{A}) = 1 - P(A)$ En effet : $A \cup \overline{A} = \Omega$ et $A \cap \overline{A} = \emptyset$ donc $P(A) + P(\overline{A}) = P(\Omega) = 1$
- $\forall (A, B) \in \mathcal{T}^2, \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
En effet : $P(A \cup B) = P(A \cup (\overline{A} \cap B)) = P(A) + P(\overline{A} \cap B) = P(A) + P(B) - P(A \cap B)$
- Pour tout $(A, B) \in \mathcal{T}^2$, si $A \subset B$ alors $P(A \setminus B) = P(B) - P(A)$.
- Pour tout $(A, B) \in \mathcal{T}^2$, si $A \subset B$ alors $P(A) \leq P(B)$.
En effet : $P(B) = P(A \cap B) + P(\overline{A} \cap B) = P(A) + P(\overline{A} \cap B) \geq 0$
- Quel que soit $(A_k)_{1 \leq k \leq n}$ une suite de n événements, $P\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k)$

Proposition. (*complément, à savoir redémontrer*)

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements,

- ❶ si $\lim_{N \rightarrow +\infty} P\left(\bigcup_{n=0}^N A_n\right) = 1$ alors : $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1$
- ❷ si $\lim_{N \rightarrow +\infty} P\left(\bigcap_{n=0}^N A_n\right) = 0$ alors : $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = 0$

Démonstration : (*faite au tableau*)

- ❶ Pour tout $n \in \mathbb{N}$, $\bigcap_{k=0}^{+\infty} A_k \subset \bigcap_{k=0}^n A_k$ donc $P\left(\bigcap_{k=0}^{+\infty} A_k\right) \leq P\left(\bigcap_{k=0}^n A_k\right)$
on a $\forall n \in \mathbb{N}, 0 \leq P\left(\bigcap_{k=0}^{+\infty} A_k\right) \leq P\left(\bigcap_{k=0}^n A_k\right)$ et $\lim_{n \rightarrow +\infty} \left(P\left(\bigcap_{k=0}^n A_k\right)\right) = 0$
donc (*passage à la limite*) $\boxed{P\left(\bigcap_{n=0}^{+\infty} A_n\right) = 0}$
- ❷ Pour tout $n \in \mathbb{N}$, $\bigcup_{k=0}^n A_k \subset \bigcup_{k=0}^{+\infty} A_k$ donc $P\left(\bigcup_{k=0}^n A_k\right) \leq P\left(\bigcup_{k=0}^{+\infty} A_k\right)$
on a $\forall n \in \mathbb{N}, P\left(\bigcup_{k=0}^n A_k\right) \leq P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leq 1$ et $\lim_{n \rightarrow +\infty} \left(P\left(\bigcup_{k=0}^n A_k\right)\right) = 1$
donc (*passage à la limite*) $\boxed{P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1}$

Proposition.

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements,

si $(A_n)_{n \in \mathbb{N}}$ est un système complet d'événements alors : $\sum_{n=0}^{+\infty} P(A_n) = 1$

Remarques en lien avec le cours sur les séries :

- Si (A_n) est une suite d'événements deux à deux incompatibles, la définition d'une probabilité de première année suffit pour montrer que la série $\sum_{n \geq 0} P(A_n)$ converge.
- Les séries convergentes sont ici toutes absolument convergentes, l'ordre de sommation n'a pas d'importance, on pourra sans ambiguïté, pour $I \subset \mathbb{N}$ utiliser la notation $\sum_{i \in I} P(A_i)$ pour les sommes.

1.3 Négligeable, presque sûr, quasi-complet.

Définitions.

Soit A un événement,

dire que A est **négligeable** signifie que $P(A) = 0$.

dire que A est **presque sûr** signifie que $P(A) = 1$.

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements,

dire que $(A_n)_{n \in \mathbb{N}}$ est un système **quasi-complet** d'événements signifie que :

- ❶ $\forall (i, j) \in \mathbb{N}^2, i \neq j \implies A_i \cap A_j = \emptyset$

- ❷ $\sum_{n=0}^{+\infty} P(A_n) = 1$ (*La série converge et sa somme vaut 1*) ou $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1$

Remarque : Si un système est complet alors il est quasi-complet. *La réciproque n'est pas toujours vraie.*

Propriétés.

Soit A un événement,

❶ Si A est quasi-impossible alors :

$$\text{pour tout événement } B, \quad P(A \cap B) = 0 \quad \text{et} \quad P(A \cup B) = P(B)$$

❷ Si A est quasi-certain alors :

$$\text{pour tout événement } B, \quad P(A \cap B) = P(B) \quad \text{et} \quad P(A \cup B) = 1$$

Démonstrations : (non faites en classe)

❶ On suppose $P(A) = 0$,

On a $A \cap B \subset A$ donc $0 \leq P(A \cap B) \leq P(A)$ donc $P(A \cap B) = 0$

et $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0 + P(B) - 0$ donc $P(A \cup B) = P(B)$

❷ On suppose $P(A) = 1$,

On a $A \subset A \cup B$ donc $P(A) \leq P(A \cup B) \leq 1$ donc $P(A \cup B) = 1$

et $P(A \cap B) = P(A) + P(B) - P(A \cup B) = P(A) + P(B) - P(B)$ donc $P(A \cap B) = P(B)$

1.4 Formule des probabilités totales (Version 1)

❶ Si $(A_k)_{1 \leq k \leq n}$ est un système complet d'événements alors pour tout événement B on a :

$$P(B) = \sum_{k=1}^n P(A_k \cap B)$$

❷ Si $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet d'événements alors pour tout événement B on a :

$$\left(\text{la série } \sum_{n \geq 0} P(A_n \cap B) \text{ est convergente et} \right) \quad P(B) = \sum_{n=0}^{+\infty} P(A_n \cap B)$$

Remarques :

• Le théorème est aussi vrai avec un système complet.

• On sait que lorsque une suite d'événements deux à deux incompatibles alors la série $\sum_{n \geq 0} P(A_n)$ converge.

Ce qui permet dans le théorème précédent d'oublier de préciser : $\sum_{n \geq 0} P(A_n \cap B)$ est convergente.

Démonstrations : (non faites en classe)

❶ $(A_k)_{1 \leq k \leq n}$ est un système complet d'événements et B est un événement :

$$\begin{aligned} P(B) &= P(B \cap \Omega) \\ &= P\left(B \cap \left(\bigcup_{k=1}^n A_k\right)\right) && \text{car } (A_k)_{1 \leq k \leq n} \text{ est un système complet} \\ &= P\left(\bigcup_{k=1}^n (A_k \cap B)\right) \\ &= \sum_{k=1}^n P(A_k \cap B) && \text{car les } (A_k \cap B) \text{ sont 2 à 2 incompatibles (\sigma-additivité).} \end{aligned}$$

❷ $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet d'événements et B est un événement :

$$\begin{aligned} P(B) &= P\left(B \cap \left(\bigcup_{n=0}^{+\infty} A_n\right)\right) && \text{car } P\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1 \text{ (voir la proposition précédente).} \\ &= P\left(\bigcup_{n=0}^{+\infty} (A_n \cap B)\right) \\ &= \sum_{n=0}^{+\infty} P(A_n \cap B) && \text{car les } (A_n \cap B) \text{ sont 2 à 2 incompatibles. (\sigma-additivité)} \end{aligned}$$

Conditionnement et indépendance.

On se place dans un espace probabilisé (Ω, \mathcal{F}, P)

2.1 Définition.

Définition.

Soient A et B deux événements tels que $P(A) \neq 0$,

On appelle probabilité de B sachant A , notée $P_A(B)$, le réel : $P_A(B) = \frac{P(A \cap B)}{P(A)}$

Théorème.

Quel que soit $A \in \mathcal{F}$ tel que $P(A) \neq 0$, l'application P_A est une probabilité sur (Ω, \mathcal{F}) .

Démonstration :

Il s'agit ici de démontrer que ① Pour tout événement B , $0 \leq P_A(B) \leq 1$, ② $P_A(\Omega) = 1$ et ③ lorsque (B_n) est une suite d'événements deux à deux incompatibles alors $P_A\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} P_A(B_n)$

2.2 Formule des probabilités composées

Théorème : (*Formule des probabilités composées*)

Soient n un entier supérieur à 2 et $(A_k)_{1 \leq k \leq n}$ une liste d'événements vérifiant $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$,

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \times P_{A_1}(A_2) \times P_{A_1 \cap A_2}(A_3) \times \dots \times P_{A_1 \cap \dots \cap A_{n-1}}(A_n)$$

2.3 Formule des probabilités totales (Version 2)

Théorème :

① Si $(A_k)_{1 \leq k \leq n}$ est un système complet d'événements non négligeables pour P (i.e : pour tout k , $P(A_k) \neq 0$) alors pour tout événement B on a : $P(B) = \sum_{k=1}^n P(A_k) \cdot P_{A_k}(B)$

② Si $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet d'événements non négligeables pour P (i.e : pour tout n , $P(A_n) \neq 0$) alors pour tout événement B on a :

$$\left(\text{la série } \sum_{n \geq 0} P(A_n) \cdot P_{A_n}(B) \text{ est convergente et} \right) \quad P(B) = \sum_{n=0}^{+\infty} P(A_n) \cdot P_{A_n}(B)$$

Démonstration : (Immédiate à partir de la version 1)

Remarques :

- Ce théorème est encore vrai avec un système complet ou un système quasi-complet d'événements.
- On pourra dans la version ② ne pas vérifier $P(A_n) \neq 0$ et quand $P(A_n) = 0$, on pose $P(A_n) \cdot P_{A_n}(B) = 0$.
- Ici aussi on peut oublier de préciser que la série converge.
- Vous pouvez ne pas écrire "non négligeables pour P (i.e : pour tout n , $P(A_n) \neq 0$)"

2.4 Formule de Bayes

Proposition :

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements,

Si $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet d'événements (tel que : pour tout n , $P(A_n) \neq 0$)

alors pour tout événement B (tel que $P(B) \neq 0$) et pour tout $k \in \mathbb{N}$ on a : $P_B(A_k) = \frac{P(A_k) \cdot P_{A_k}(B)}{\sum_{n=0}^{+\infty} P(A_n) \cdot P_{A_n}(B)}$

Remarques :

- En pratique il est souvent plus simple de redémontrer cette formule.
- Ce théorème est vrai pour un système complet et pour un système quasi-complet d'événements.
- Ce théorème est vrai avec un système complet d'événements fini. (*Théorème vu en première année*).

2.5 Indépendance

Définition :

Soient A et B deux événements,

Dire que A et B sont **indépendants** signifie que : $P(A \cap B) = P(A)P(B)$

Propositions :

Soient A et B deux événements,

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

❷ Si A et B sont indépendants alors A et \bar{B} sont indépendants, \bar{A} et B sont indépendants et \bar{A} et \bar{B} sont indépendants.

Démonstration :

Remarque : Dire que (A_n) est une suite d'événements **2 à 2 indépendants** signifie que

$$\forall (i, j) \in I^2, \quad i \neq j \implies P(A_i \cap A_j) = P(A_i)P(A_j)$$

Définition :

Soit $(A_n)_{n \in I}$ une suite finie ou dénombrable d'événements.

Dire que les A_n sont **mutuellement indépendants** signifie que :

$$\text{pour toute partie finie } J \text{ de } I, \quad P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

Autrement dit :

Pour tout entier $k \geq 2$, et k indices $i_1 < \dots < i_k$ on a : $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times \dots \times P(A_{i_k})$.

Remarque : si les A_n sont mutuellement indépendants alors les A_n sont 2 à 2 indépendants,

la réciproque n'est pas toujours vraie.

Proposition :

Soient $(A_n)_{n \in I}$ une suite finie ou dénombrable d'événements.

et $(B_n)_{n \in I}$ une suite d'événements vérifiant : $\forall n \in I, B_n = A_n$ ou $B_n = \bar{A}_n$

❶ Si les A_n sont 2 à 2 indépendants alors les B_n sont 2 à 2 indépendants.

❷ Si les A_n sont mutuellement indépendants alors les B_n sont mutuellement indépendants.

Autrement dit : Si dans une suite d'événements deux à deux (resp. mutuellement) indépendants on remplace un ou plusieurs événements par leur contraire, on obtient toujours une suite d'événements deux à deux (resp. mutuellement) indépendants.

Remarque : Vous pouvez ne pas écrire cette proposition.

Variables aléatoires réelles.

On considère ici un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

3.1 Définition.

Définition

Une variable aléatoire réelle est une application de Ω dans \mathbb{R} vérifiant :
 Pour tout $a \in \mathbb{R}$, $[X \leq a]$ est un événement.

Remarques :

- Comme pour les tribus vous ne devriez pas être interrogés sur cette définition.
- On note $[X \leq a]$ l'événement $\{\omega \in \Omega \mid X(\omega) \leq a\}$
- L'ensemble des valeurs prises par X est $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$
- Pour tout $(a, b) \in \mathbb{R}^2$, $[X \geq a]$, $[a \leq X \leq b]$ sont des événements.

Proposition : Pour tout I intervalle de \mathbb{R} , $[X \in I]$ est un événement.

3.2 Fonction de répartition.

Définition

Soit X une variable aléatoire de $(\Omega, \mathcal{F}, \mathbb{P})$,
 On appelle fonction de répartition de X la fonction $F : \mathbb{R} \rightarrow [0, 1]$
 $x \mapsto \mathbb{P}([X \leq x])$

Remarques :

- Quand plusieurs variables aléatoires sont présentes on note $F_X : \mathbb{R} \rightarrow [0, 1]$
 $x \mapsto \mathbb{P}([X \leq x])$
- C'est à rapprocher de la notion de fréquences cumulées en statistiques.

Proposition :

Pour tout $(a, b) \in \mathbb{R}^2$ tel que $a \leq b$ on a :

$$\mathbb{P}(X \in]a, b]) = F(b) - F(a)$$

Démonstration.

Proposition :

Soit X une variable aléatoire de fonction de répartition F alors :

- ① F est une fonction croissante sur \mathbb{R} .
- ② F est partout continue à droite.
- ③ $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$

Démonstration :

3.3 Indépendance de deux variables aléatoires.

Définition (*définition du programme*)

Soient X et Y deux variables aléatoires sur le même espace probabilisé (Ω, \mathcal{F}, P) ,
 Dire que X et Y sont indépendantes signifie que :
 quel que soit le couple (I, J) d'intervalles de \mathbb{R} , $P((X \in I) \cap (Y \in J)) = P(X \in I) \times P(Y \in J)$

Théorème (*caractérisation*)

Soient X et Y deux variables aléatoires de (Ω, \mathcal{F}, P) ,
 X et Y sont indépendantes si, et seulement si, $\forall (x, y) \in \mathbb{R}^2$, $P((X \leq x) \cap (Y \leq y)) = P(X \leq x) \times P(Y \leq y)$

3.4 Indépendance de n variables aléatoires.

Définition

Soient n un entier supérieur ou égal à 2 et $(X_k)_{1 \leq k \leq n}$ une liste de variables aléatoires de (Ω, \mathcal{F}, P) ,
 Dire que les X_k sont (*mutuellement*) indépendantes signifie que :
 quelle que soit la liste $(I_k)_{1 \leq k \leq n}$ d'intervalles de \mathbb{R} , $P\left(\bigcap_{k=1}^n (X_k \in I_k)\right) = \prod_{k=1}^n P(X_k \in I_k)$

Théorème (*caractérisation*)

Soient n un entier supérieur ou égal à 2 et $(X_k)_{1 \leq k \leq n}$ une liste de variables aléatoires sur le même espace probabilisé (Ω, \mathcal{F}, P) ,
 Les X_k sont (*mutuellement*) indépendantes si, et seulement si, :
 quelle que soit la liste $(x_k)_{1 \leq k \leq n}$ de réels, $P\left(\bigcap_{k=1}^n (X_k \leq x_k)\right) = \prod_{k=1}^n P(X_k \leq x_k)$

Remarque :

On parle aussi de "liste mutuellement indépendante" (*pour "liste de variables aléatoires mutuellement indépendantes"*).

3.5 Indépendance d'une suite de variables aléatoires.

Définition

Soit $(X_n)_{n \in \mathbb{N}}$ une liste de variables aléatoires de (Ω, \mathcal{F}, P) ,
 Dire que les variables de $(X_n)_{n \in \mathbb{N}}$ sont (*mutuellement*) indépendantes signifie que :
 toute liste finie extraite de la suite est (*mutuellement*) indépendante.

3.6 Propriétés de l'indépendance mutuelle.

Théorème :

Soit $(X_k)_{1 \leq k \leq n}$ une liste de variables aléatoires de (Ω, \mathcal{F}, P) ,

❶ Si X_1, \dots, X_n sont (*mutuellement*) indépendantes **alors** toute sous-famille de (X_1, \dots, X_n) l'est aussi.

❷ (Lemme des coalitions) Soient $f : \mathbb{R}^p \rightarrow \mathbb{R}$ et $g : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$ deux fonctions, si $X_1, \dots, X_p, \dots, X_n$ sont (*mutuellement*) indépendantes **alors** $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

❸ Soient f_1, \dots, f_n des fonctions de \mathbb{R} dans \mathbb{R} , si X_1, \dots, X_n sont (*mutuellement*) indépendantes **alors** $f_1(X_1), \dots, f_n(X_n)$ sont (*mutuellement*) indépendantes.