Théorie basique des probabilités

Notations Rappels

- $I \subset \mathbf{N}$: partie finie ou dénombrable.
- Ainsi $\bigcup_{i\in I}$, $\bigcap_{i\in I}$, $\sum_{i\in I}$ sont des notations signifiant suivant le contexte réunion/intersection/somme [a] finie ou dénombrable.
- Ω : ensemble au plus dénombrable.
- $(\Omega, \mathscr{T}, P)$: espace probabilisé.
- $\mathscr{T} = \mathscr{P}(\Omega)$ si Ω est fini, sinon une tribu adaptée à l'expérience analysée.

a. Si I est fini, c'est une somme simple, sinon c'est la limite (somme) d'une série convergente.

I Vocabulaire

A) Dictionnaire probas/théorie des ensembles (bilingue)

Notion	Ensembles	Probas
Ω	partie pleine	univers
$\varnothing$	ensemble vide	évènement impossible
$\omega \in \Omega$	élément	observation
$A \in \mathscr{T}$	sous-ensemble	évènement
$\omega \in A$	appartenance	ω est une réalisation de A
$A \subset B$	inclusion	A implique B
$\bar{A}$	ensemble complémentaire	évènement contraire
$A \cup B$	réunion	évènement «A ou B»
$A \cap B$	intersection	évènement «A et B»
$A \cap B = \varnothing$	ensembles disjoints	évènements
•$\bigcup_{i\in I} A_i = \Omega$ •$i \neq j \Leftrightarrow A_i \cap A_j = \varnothing$	partition de Ω	système complet d'évènements (SCE)
•$P\left(\bigcup_{i\in I} A_i\right) = 1$ •$i \neq j \Leftrightarrow A_i \cap A_j = \varnothing$	Notions d'analyse et pas de théorie des ensembles puisqu'elles font intervenir la fonction P !	Système quasi-complet d'évènements
$P(A) = 1$		Évènement quasi-certain
$P(A) = 0$		Évènement quasi impossible

T₁> Ce lexique est fondamental : il permet de passer de la formulation en français d'un énoncé à sa traduction mathématique et *vice-versa*.

B) Aléatoire

Expérience : processus reproductible aboutissant à un résultat observable ω.

Expérience aléatoire : le résultat observable est variable lors des réalisations de l'expérience : ω n'est pas constant.

Stochastique : relatif à la théorie des probabilités.

Univers : Typ. Ensemble contenant toutes les valeurs possibles de ω. Son existence est postulée.

C) Probabilité : Attention au sens de ce mot

Différentes acceptions du terme :

a. Typ. Réel Probabilité = Valeur de $P(A)$.

b. Typ. Fonction Probabilité = La fonction définie sur $\mathscr{T}$ (notée P en général).

c. Typ. Fonction Probabilité = Mesure de probabilité au sens **b**.

II Mesure de probabilité

A) Propriétés

a. *Normalisation.* : $P(\Omega) = 1$

b. *Positivité.* $\forall A \in \mathscr{T} \quad P(A) \geq 0$.

c. *σ-additivité (s'appelle additivité finie si I est fini)* . Si $(A_i)_{i\in I}$ est une famille d'évènements deux à deux incompatibles, la série de t.g $P(A_i)$ converge et :
$$P\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} P(A_i)$$

Rem.1 | La convergence de la série n'est pas à vérifier : elle est exigée dans la définition.

Conséquences de ces propriétés :

a. $P(\bar{A}) = 1 - P(A)$.

b. $P(\varnothing) = 0$.

c. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

d. $A \subset B \Rightarrow P(A) \leq P(B)$.

B) Construction de probabilités

[Distribution de masse sur Ω] Déf.1

Toute fonction π définie sur $\Omega := \{\omega_i \quad i \in I\}$ vérifiant :

a. $\forall i \in I \quad p_i := \pi(\omega_i) \geq 0$

b. La série de t.g p_i converge vers 1 : $\sum_{i\in I} p_i = 1$

[Construction des mesures de proba] Thm.1

La donnée d'une fonction de masse π sur Ω définit une mesure de probabilité P sur Ω par :
$$\forall A \in \mathscr{T} \quad P(A) = \sum_{\substack{i\in I \\ i \text{ tq.} \omega_i \in A}} \pi(\omega_i)$$

NB : *S'applique au cas des univers au plus dénombrables.*

Typ. On écrit $P(\{\omega\})$, mais $\pi(\omega)$.

T₂>

a. En pratique pour montrer qu'on a probabilisé $(\Omega, \mathscr{T})$, on vérifie que la fonction donnée est une fonction de masse.

b. Dans le cas où Ω est non dénombrable (typiquement : un intervalle de $\mathbf{R}$), on recourt au même procédé, la fonction de masse étant remplacée par une densité de probabilité.

Rem.2 | Si l'ensemble I (et donc Ω) est fini, en prenant π constante égale à $\#\Omega$, on obtient comme mesure P la mesure appelée probabilité uniforme ou équiprobabilité sur Ω. Sinon, la condition« π est constante » est incompatible avec la condition **b.** de **déf. 1** : on ne peut équiprobabiliser Ω.

III Probabilité conditionnelle

a) Mesure conditionnelle

[mesure conditionnelle sous l'hypothèse H] **Déf.2**

Si $H \in \mathscr{T}$ est tel que $P(H) \neq 0$, la fonction notée P_H définie sur $\mathscr{T}$ par :

$$A \mapsto P_H(A) := \frac{P(A \cap H)}{P(H)}$$

est une mesure de probabilité sur Ω.

Typ. Fonction La probabilité conditionnelle est une probabilité au sens de I c) b.

T3> La fonction P_H vérifie donc toutes les propriétés énoncées en II A). En particulier, $P_H(\bar{A}) = 1 - P_H(A)$.

Rem.3

- **a.** La fonction P_Ω est simplement la fonction P.
- **b.** $P(A)$ s'appelle probabilité de A *a priori* (c'est-à-dire dans l'état des connaissances à l'issue de l'expérience/observation) et $P_H(A)$ s'appelle probabilité de A *a posteriori* puisque, les connaissances ont évolué dans la mesure où l'on *sait* que H a eu lieu.

Rem.4 Par définition, et même si $P(H) = 0$

$$P(A \cap H) = P_H(A) \times P(H)$$

b) Formules usuelles

[Formule des probas composées] **Thm.2**

Soit $A_1, \dots, A_n$ n évènements. On pose $\hat{A}_0 = \Omega$ et $\hat{A}_k = \bigcap_{j=1}^{k} A_j$ pour tout $k \in \{1 \dots n\}$.

Alors : $P(\hat{A}_n) = \prod_{k=1}^{n} P(A_k | \hat{A}_{k-1})$.

T4> Formule particulièrement adaptée à une analyse chronologique de l'expérience, ou pour calculer la probabilité d'un évènement réalisé par un seul chemin sur l'arbre.

[Formule des proabs totales] **Thm.3**

Si $(H_i)_{i \in I}$ est un SCE ou un SQCE de Ω, alors la série de t.g. $P(A|H_i)P(H_i)$ converge et on a

$$P(A) = \sum_{i \in I} P(A|H_i)P(H_i)$$

T5>

- **a.** Formule adaptée pour conditionner suivant les cas possibles, ou pour calculer la probabilité d'un évènement apparaissant plusieurs fois sur l'arbre
- **b.** Cas courant : SCE binaire : $(H, \bar{H})$ et $P(A) = P(A|H)P(H) + P(A|\bar{H})P(\bar{H})$

.

[Formule de Bayes] **Thm.4**

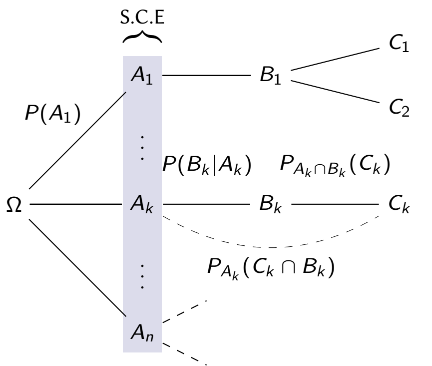
Si A et H sont deux évènements de probabilité non nulle :

$$P(H|A) = \frac{P(A|H)}{P(A)} \times P(H)$$

T6> Formule permettant de renverser causes et effets dans le calcul stochastique. Le dénominateur $P(A)$ se calcule dans la plupart des cas par probabilités totales.

c) Arbres : mode d'emploi

$\star, \bullet, \circ$: des évènements



- **a.** **Chemin :** le chemin $[\Omega - \bullet - \star - \circ \dots]$ sur l'arbre représente l'évènement $[\bullet \cap \star \cap \circ \cap \dots]$.
- **b.** **Probabilité conditionnelle :** le poids d'un chemin de racine $\bullet$, mettons : $\bullet - \circ$ est égal à la probabilité conditionnelle $P_\bullet(\circ)$ (En particulier : si $\bullet = \Omega$, on calcule les probabilités *a priori*, voir **Rem.4**).
- **c.** **Formule des probabilités composées :** les poids se multiplient le long des branches (composition des poids).
- **d.** **Systèmes complets :** tout niveau d'arborescence est un S.C.E (ou un S.Q.C.E).
- **e.** **Formule des probabilités totales :** si l'évènement $\bullet$ apparaît plusieurs fois *sur un même niveau* d'arboresence, $P(\bullet)$ est la somme des probabilités des occurrences de $\bullet$ sur l'arbre.
- **f.** ☠ Un évènement ne peut jamais apparaître à deux niveaux d'arborescence distincts. Exemple : suite de pile ou face.

IV Indépendance

a) Cas de deux évènements

Deux évènements A et B sont $(P-)$indépendants si et seulement si l'une des conditions suivantes équivalentes est vraie :

- **a.** $P(A \cap B) = P(A) \times P(B)$
- **b.** $P(A|B) = P(A)$
- **c.** $P(B|A) = P(B)$

Rem.5 En général, on devrait parler de $P-$indépendance, car cette propriété dépend fortement de la fonction P de probabilité avec laquelle on évalue les nombres apparaissant dans **a**-**c**.

b) Cas de $n \geq 2$ évènements

Cette définition étend la précédente. Les évènements $A_1, \dots A_n$ sont mutuellement indépendants si et seument si :

C2) $\forall i, j$ tels que $i < j$, A_i et A_j sont indépendants.

C3) $\forall i, j, k$ tels que $i < j < k$:

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$$

$\vdots$

Ck) $\forall i_1, \dots, i_k$ tels que $i_1 < \dots < i_k$:

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times \dots \times P(A_{i_k})$$

$\vdots$

Cn) $P(A_1 \cap \dots A_n) = P(A_1) \times \dots \times P(A_n)$.

Rem.6 Dans **Cj**), il y a $\binom{n}{j}$ conditions à vérifier. Il y a au total $2^n - n - 1$ conditions à vérifier.

Rem.7 La mutuelle indépendance est rarement à prouver : elle est soit une hypothèse, soit une conséquence du modèle.

c) Cas d'une famille quelconque d'évènements

La famille $\mathscr{F}$ est une famille d'évènements mutuellement indépendants si toute sous-famille finie d'évènements extraite de $\mathscr{F}$est une famille d'évènements mutuellement indépendants.

d) Propriétés

- **a.** Permet de transformer les probas d'intersections en produit de probas dans les calculs.
- **b.** La propriété de mutuelle indépendance des évènements d'une famille est préservée en remplaçant dans la famille autant d'évènements que souhaité par leur contraire.

T7> Penser à citer ces propriétés dans les calculs.