CH9 – Théorie basique des probabilités

Plan du chapitre

Liste des définitions

Liste des techniques de base

Grille d'analyse des exercices

Exercice	Question	T	Référence(s)	Commentaires/remarques

1. **T_0** : technique ancestrale. Pas listée dans les techniques de base.
2. **Déf** : pas de technique livrée. Revenir à la définition.
3. **C** : utilisation d'un résultat de cours (théorème, proposition, etc.)
4. $\boxed{\star}$ Question discriminante et plus difficile : demande raisonnement et enchaînement de techniques.

Faire attention à ce qui suit

Dénombrabilité

■ Définition 1 .... [Ensemble infini et dénombrable]
Ensemble en bijection avec $\mathbf{N}$.

■ Remarque 1.
Intuitivement, c'est un ensemble dont on peut numéroter les éléments.

■ Exemple 1.
Les ensembles $\mathbf{N}$, $\mathbf{N}^\star, \mathbf{Z}, \mathbf{Q}$ sont infinis et dénombrables. L'intervalle $[0,1]$ ne l'est pas.

■ Définition 2 .... [Ensemble au plus dénombrable]

Ensemble qui est fini ou alors infini et dénombrable.

■ Remarque 2.
Ne pas confondre ensemble fini et ensemble borné. : [0,1] est borné mais infini et non dénombrable

Opérations ensemblistes, sommes et dénombrabilité

Si I est un ensemble au plus dénombrable, $\bigcup_{i\in I}, \bigcap_{i\in I}, \sum_{i\in I}$ sont des notations signifiant suivant le contexte
Ensemble d'indices celui-là inspire méfiance
réunion/intersection/somme [1] finie ou dénombrable.

1 Vocabulaire des probabilités

La théorie des probabilités fournit un cadre mathématique permettant d'analyser mathématiquement des *expériences aléatoires*, en mettant en évidence des comportements réguliers sur des phénomènes considérés comme imprévisibles.

A) Ce qu'est une expérience aléatoire

■ Définition 3 (vraie vie / réalité) .... [Expérience]
Processus reproductible aboutissant à un résultat observable ω.

■ Définition 4 .... [Expérience alétoire]
le résultat observable de l'expérience est variable lors de réalisations de l'expérience : ω n'est pas unique.

■ Définition 5 .... [Univers associé à une expérience]
Ensemble Ω contenant toutes les valeurs possibles de ω.
fait le trait d'union entre vraie vie et maths.

■ Remarque 3.

1. On admet que pour toute expérience, il existe un univers décrivant les observations possibles à son issue.
2. Une expérience aléatoire a donc par définition un univers contenant au moins deux éléments. Sinon, l'expérience est dite déterministe.

■ Définition 6 .... [Observation - Évènement élémentaire]

Si ω est un élément de Ω, ω s'appelle observation, et $\{\omega\}$ s'appelle évènement élémentaire.
TYP sous-ensemble de Ω

1. Si I est fini, c'est une somme contenant un nombre fini de réels, sinon c'est la limite d'une série convergente, c'est-à-dire sa somme.

■ **Définition 7** .... [Stochastique]

Relatif à la théorie des probabilités.

B) Dictionnaire bilignue probas/théorie des ensembles

But : formuler des évènements

■ **Définition 8** .. [évènement, certain, impossible, implique, contraire, incompatibles, SCE]

Notation	Terme ensembliste	Traduction probabiliste
$\mathscr{T}$	Tribu ($\mathscr{T} = \mathscr{P}(\Omega)$ en sup.)	Ensemble des évènements
Ω	partie pleine	univers
$\varnothing$	ensemble vide / partie vide	évènement impossible
$\omega \in \Omega$ ω atome	élément de Ω	ω est une observation
$A \in \mathscr{T}$	élément de $\mathscr{T}$	A est un évènement
$\omega \in A$	ω appartient à A	ω est une réalisation de l'évènement A = sous-ensemble de Ω
$A \subset B$	A est inclus dans B	A implique B
$\bar{A}$	ensemble complémentaire de A	évènement contraire de A
$A \cup B$	réunion	évènement [A ou B]
$A \cap B$	intersection	évènement [A et B]
$A \cap B = \varnothing$	A et B sont disjoints	A et B sont incompatibles
$(A_i)_{i\in I}$ t.q. $\bullet \bigcup_{i\in I} A_i = \Omega$ $\bullet i \neq j \Leftrightarrow A_i \cap A_j = \varnothing$	$(A_i)_{i\in I}$ est une partition de Ω	$(A_i)_{i\in I}$ est un système complet d'évènements (SCE)

vous pouvez les penser comme cela et oublier $\mathscr{T}$

T1 **Usage du dictionnaire**

Cette technique est de loin la technique la plus importante de la théorie des probabilités. Le dictionnaire permet de passer de la formulation en français d'un énoncé à sa traduction mathématique et *vice-versa*. Il faut donc toujours commencer par :

1. Introduire des évènements correctement définis et les plus élémentaires possibles.. Par exemple : P_k : «on a eu pile à l'issue du k-ème lancer» est un évènement d'intérêt et bien défini. En revanche : P : «on a pile» n'est pas bien défini.
2. Utiliser le dictionnaire et les évènements définis dans **1.** pour exprimer les évènements plus élaborés à l'aide des opérations ensemblistes.

Sans formation correcte des évènements, impossible de calculer des probabilités.

■ Exemple 2.

On lance un dé. On observe le score. (analysable en sup)

1. Lister les observations possibles.
2. En déduire l'univers Ω.
3. Expliciter les évènements suivants et dire si ils sont élémentaires :
   - **a)** A_2 :« le score obtenu est plus grand que 2».
   - **b)** A_3 :«le score obtenu est plus grand que 3 ».
   - **c)** Π :«le score obtenu est pair».
   - **d)** J :«le score obtenu est impair».
   - **e)** D : «On obtient 2».
4. Écrire en langage probabiliste :
   - **a)** [Observer] un score plus grand que 3 implique [d'observer] un score plus grand que 2
   - **b)** Observer un score impair et observer un score pair sont incompatibles.
   - **c)** [Observer] 2 est une réalisation de [l'évènement « observer] un score pair. »

✓ : élém ✗ : pas élém

1 - $\omega = 1$, ou $\omega = 2$, ..., $\omega = 6$

2 $\Omega \underset{\text{déf}}{=} \{1, 2, \dots, 6\}$

déf 8 → Ce sont des sous-ensembles de Ω

3. a) $A_2 = \{3, 4, 5, 6\}$ ✗ (en français : plus grand = strict plus grand)

b) $A_3 = \{4, 5, 6\}$ ✗

c) $\Pi = \{2, 4, 6\}$ ✗ extension

d) $J = \{1, 3, 5\}$ ✗ Rem : $\overline{\Pi} = J$

e) $D = \{2\}$ ✓

4-a) $A_3 \subset A_2$ param

4-b) $\Pi \cap J = \emptyset$

4-c) $2 \in \Pi$

■ Exemple 3.

On dispose d'une pièce équilibrée, on la lance une infinité de fois. On observe la suite des lancers.

1. Définir l'univers associé à cette expérience.
2. Pout tout entier $k \geq 1$, on note F_k l'évènement : le k-ème lancer donne face. Est-ce que F_1 est un évènement élémentaire ?
3. Exprimer à l'aide des évènements F_k les évènements suivants :
   - **a)** A_1 : «Les 5 premiers lancers ont donné face».
   - **b)** A_2 : «Le premier face a eu lieu au lancer 14».
   - **c)** A_3 : «Sur les 78 premiers lancers, il y a eu un face (au moins)» (*passer par le contraire*)
4. Soit $n \geq 2$.
   - **a)** B_n : «les n premiers lancers ont donné face».
   - **b)** C_n : « Le premier pile a eu lieu avant le lancer n (inclus)».
   - **c)** D_n : «Le premier pile a eu lieu au lancer n» (à l'aide des C_k).
   - **d)** E_n : «Le deuxième pile a eu lieu au lancer n».
5. **a)** G : «on n'a jamais observé pile».
   b) F : «il y a eu au moins un face ».
6. Soit $n \geq 1$. Traduire : «N'avoir observé aucun pile implique que les n premiers lancers ont donné face.»

très souvent (≃100% des exos)

2 spécifique à la spé : pas analysable en sup

déf 8 → sous-ensemble de Ω

1- Ici une observation = une suite de lancers

donc Ω = un ensemble de suites

$\Omega = \left\{ (x_n)_{n \in \mathbb{N}^*} \quad \begin{matrix} x_1 \in \{\text{pile}, \text{face}\} \\ x_2 \in \{\text{pile}, \text{face}\} \\ \vdots \end{matrix} \right\}$

2 - F_1 n'est pas un évènement élémentaire car : $x = (\text{face}, \text{face}, \dots\dots) \in F_1$ et $y = (\text{face}, \text{pile}, \text{pile}, \text{pile}, \dots) \in F_1$ donc F_1 n'est pas un singleton.

rem : $F_1 = \{(x_n)_{n \in \mathbb{N}^*} \in \Omega \text{ tq } x_1 = \text{face}\}$ [poly]

lois de Morgan

$$\boxed{\begin{aligned} \overline{A \cup B} &= \overline{A} \cap \overline{B} \\ \overline{A \cap B} &= \overline{A} \cup \overline{B} \\ \overline{\overline{A}} &= A \end{aligned}}$$

3-a) $A_1 = F_1 \cap F_2 \cap F_3 \cap F_4 \cap F_5$

b) $A_2 = \overline{F_1} \cap \dots \cap \overline{F_{13}} \cap F_{14}$

c) D'après l'indication, on calcule

$\overline{A_3}$ = " les 78 premiers lancers ont donné pile "

$= \overline{F_1} \cap \dots \cap \overline{F_{78}}$

donc :

$A_3 = \overline{\overline{A_3}} = \overline{\overline{F_1} \cap \overline{F_2} \cap \dots \cap \overline{F_{78}}}$

$$\boxed{(\star)\ A_3 \underset{\text{lois de Morgan}}{=} F_1 \cup F_2 \cup \dots \cup F_{78}}$$

rem : vous n'êtes pas obligés de passer par $\overline{A_3}$ si $(\star)$ est évident

4-a) $B_n = \bigcap_{k=1}^{n} F_k$

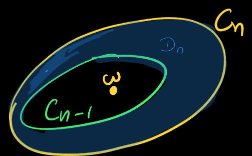
b) C_n = " sur les n 1ers lancers, il y a eu au moins un pile "

$C_n = \bigcup_{k=1}^{n} \overline{F_k}$

c) $D_n = C_n \cap \left(\bigcap_{k=1}^{n-1} F_k \right)$

$= C_n \cap \overline{C_{n-1}}$

rem : $(\star)\ C_n \supset C_{n-1}$ car si on a observé pile sur les $n-1$ premiers lancers, on l'a observé a fortiori sur les n premiers :



$\boxed{C_{n-1} \text{ implique } C_n}$

Comme $(\star)$ est vraie :

$D_n = C_n \setminus C_{n-1}$

5-a) $G = \bigcap_{n=1}^{+\infty} F_n$ = tous les lancers (à tout rang n) donnent face

$F = \bigcup_{n=1}^{+\infty} F_n$ car $(\omega_n)_{n \geq 1} \in F \iff \exists n \in \mathbb{N}^* \ \omega_n = \text{face}$

4-d) $E_n = \bigcup_{k=1}^{n-1} \overbrace{\boxed{\overline{F_k}} \cap \boxed{\overline{F_n}} \cap \left(\boxed{\bigcap_{\substack{j=1 \\ j \neq k}}^{n-1} F_j} \right)}^{H_k}$

H_k : sur les n 1ers lancers : j'ai eu :
- $\boxed{\text{pile au lancer } k}$
- et $\boxed{\text{pile au lancer } n}$
- et $\boxed{\text{face aux autres lancers}}$

et $E_n = H_1 \cup H_2 \cup \dots \cup H_{n-1}$.

TD : exos 78, 79 (classique), 81, 82

6- $\boxed{\forall n \geq 1 \quad G \subset B_n}$

2 Mesure de probabilité But : mesurer le hasard sur des évènements

■ Définition 9 .... [Tribu (d'évènements)]
Les tribus sont les structures garantissant que les opérations standard sur les évènements donnent encore des évènements, à savoir les opérations $\overline{}$,$\bigcup\limits_{i\in I}$, et $\bigcap\limits_{i\in I}$. Autrement dit, ce sont des ensembles d'évènements stables par les opérations précédentes.

A) Propriétés axiomatiques d'une probabilité

■ Définition 10 .... [Mesure de probabilité ou probabilité]
Soit Ω un univers, équipé d'une tribu $\mathscr{T}$. Une *probabilité sur* Ω est une fonction P définie sur $\mathscr{T}$ vérifiant les conditions suivantes :

1. *Normalisation.* : $P(\Omega) = 1$.
2. *Positivité.* $\forall A \in \mathscr{T} \quad P(A) \geq 0$.
3. *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, alors :
   a) la série de t.g $P(A_i)$ converge.
   b) De plus, il est vrai que $P\left(\bigcup\limits_{i\in I} A_i\right) = \sum\limits_{i\in I} P(A_i)$.

■ Remarque 4.
La convergence de la série dans **3.a)** n'est pas à vérifier : elle est exigée dans les propriéts de la fonction P.

■ Remarque 5.
Attention au sens du mot *probabilité :*

1. **TYP** Fonction Probabilité = La fonction définie sur $\mathscr{T}$ (notée P en général, `déf. 8`)
2. **TYP** Fonction Probabilité = synonyme de mesure de probabilité (plus court que mesure de probabilité)
3. **TYP** Réel Probabilité = probabilité d'un évènement A = réel $P(A)$ = valeur de la fonction éponyme sur un point $A \in \mathscr{T}$ de son domaine.

T₂ Comment utiliser la σ-additivité

1. On justifie que la suite $(A_n)_{n\geq 1}$ est une suite d'évènements 2 à 2 imcompatibles.
2. On ajoute ensuite : «*par σ-additivité la série de t.g. $P(A_n)$ est convergente et sa somme vaut $P\left(\bigcup\limits_{n=1}^{+\infty} A_n\right)$:*»

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

par rapport à la rem 4 : vous n'avez pas à prouver la convergence, mais vous devez la mentionner

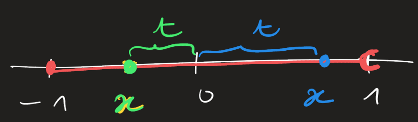
---

■ Définition 11 .... [Espace probabilisable - espace probabilisé]
Un univers Ω équipé d'une tribu d'évènements $\mathscr{T}$ est dit probabilisable. Un espace probabilisatble équipé d'une fonction de probabilité P définie sur $\mathscr{T}$ fait du triplet $(\Omega, \mathscr{T}, P)$ un espace probabilisé.

Exo75 (séries) → classique Agro. Cette série intervient en probas

But : prouver que $\forall x\in[-1,1[$, $\sum_{k=1}^{+\infty}\frac{x^k}{k}$ existe et vaut $-\ln(1-x)$.

Preuve : • $\boxed{\text{fixons } x\in[-1,1[}$ je considère ensuite t un réel entre 0 et x (ou x et 0 si $x<0$)



• Ensuite je fixe un entier $n\geq 1$ et j'écris :

$$\forall t\in[0,x] \text{ ou } [x,0] \quad \sum_{k=0}^{n-1} t^k \underset{t\neq 1}{=} \frac{1-t^n}{1-t} = \frac{1}{1-t} - \frac{t^n}{1-t}$$

J'intègre terme à terme de 0 à x par linéarité :

$$\sum_{k=0}^{n-1}\int_0^x t^k\,dt = -\ln(1-t) - \int_0^x \frac{t^n}{1-t}$$

$$\sum_{k=0}^{n-1}\frac{x^{k+1}}{k+1} = -\ln(1-x) - \int_0^x\frac{t^n}{1-t}\,dt$$

Par glissement d'indices :

$$-\ln(1-x) = \underbrace{\sum_{k=1}^{n} \frac{x^k}{k}}_{\text{somme partielle } S_n \text{ de } \sum_{k\geq 1}\frac{x^k}{k}} + \underbrace{\int_0^x\frac{t^n}{1-t}\,dt}_{I_n}$$

• Enfin, je montre que $I_n \xrightarrow[n\to+\infty]{} 0$, ce qui prouve :

- (S_n) converge.
- $\sum_{k=1}^{+\infty}\frac{x^k}{k} = -\ln(1-x)$

Par croissance de l'intégrale et inégalité triangulaire :

$$|I_n| \leq \int_0^x \frac{t^n}{1-t}\,dt \quad \text{si } x\geq 0$$

① si $x\geq 0$: $1-t \geq 1-x \geq 0$ car $t\leq x\leq 1$

donc $0\leq \frac{1}{1-t} \leq \frac{1}{1-x}$ par ↘ de $u\mapsto\frac{1}{u}$ sur $\mathbb{R}_+^*$

$$0\leq \frac{t^n}{1-t} \leq \frac{t^n}{1-x}$$

donc en primitivant :

$$\forall n\geq 1 \quad 0\leq I_n \leq \frac{x^{n+1}}{n+1}\times\frac{1}{1-x} \leq \frac{1}{n+1}\times\frac{1}{1-x}$$

d'où $I_n \xrightarrow[n\infty]{} 0$ par les gendarmes

② Si $x<0$: $0>x\geq t\geq -1$

$$|I_n| = \left|\int_x^0 \frac{t^n}{1-t}\right| \leq \int_x^0 \frac{|t|^n}{1-t}\,dt$$

donc : $-x\leq -t\leq 1$

$0\leq 1-x\leq 1-t\leq 2$

$\frac{1}{1-t}\leq\frac{1}{1-x}$ par ↘ de la fonction inverse.

Même conclusion : $|I_n| \underset{n\infty}{=} o(1)$.

q2) On a le droit de choisir $x=\frac{1}{2}$.

Rem : utiliser ce résultat dans Exo82 q2)

■ Proposition 1 .... [Propriétés d'une probabilité P]

1. Pour tout évènement A de $\mathscr{T}$: $P(\bar{A}) = 1 - P(A)$.
2. $P(\varnothing) = 0$.
3. Pour tout évènements A et B de $\mathscr{T}$:$P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Formule de Poincaré.
4. $A \subset B \Rightarrow P(A) \leq P(B)$. Croissance
   et $P(B \setminus A) = P(B) - P(A)$

■ Définition 12 .... [SQCE, quasi-certain/presque sûr, quasi-impossible]

Notation	Traduction ensembliste	Traduction probabiliste
•$P\left(\bigcup\limits_{i \in I} A_i\right) = 1$ •$i \neq j \Leftrightarrow A_i \cap A_j = \varnothing$	Notions d'analyse et pas de théorie des ensembles puisqu' elles font intervenir la fonction P !	$(A_i)_{i \in I}$ est un système quasi- complet d'évènements (SQCE)
$P(A) = 1$ ⚠ en général $A \neq \Omega$		A est un évènement quasi-certain/presque sûr
$P(A) = 0$ évènement pas forcément impossible : $A \neq \varnothing$ en gal mais A est improbable		A est un évènement quasi impossible

■ Exercice 1.

Alice et Bob lancent une pièce à tour de rôle. Alice commence. Le gagnant est le premier à faire `pile`. Les lancers sont supposés mutuellement indépendants. Soit N l'évènement : «La partie ne s'arrête jamais», et F_k l'évènement : «le k-ème lancer donne `face`». Le but est de montrer que N est quasi-impossible.

1. Soit $n \geq 1$ un entier et C_n:«Les n premiers lancers donnent face». Montrer que pour tout entier $P(N) \leq P(C_n)$.
2. En déduire que pour tout entier $n \geq 1$: $P(N) \leq \dfrac{1}{2^n}$. Conclure.

B) Construction de probabilités

■ Définition 13 .... [Distribution/fonction de masse sur Ω]

Soit $\Omega = \{\omega_i \quad i \in I\}$ un ensemble au plus dénombrable. On appelle distribution ou fonction de masse sur Ω toute fonction π définie sur Ω vérifiant :

1. $\forall i \in I \quad p_i := \pi(\omega_i) \geq 0$ (π est une fonction à valeurs positives).
2. La série de t.g p_i converge vers 1 : $\sum\limits_{i \in I} p_i = 1$.

pour tout ensemble A :

$$A = \bigcup_{a \in A} \{a\} = \{a \quad a \in A\}$$

■ Théorème 1 .... [*Toute fonction de masse induit une mesure de proba*]

La donnée d'une fonction de masse π sur Ω définit une mesure de probabilité P sur $(\Omega, \mathscr{T})$ par :

$$\forall A \in \mathscr{T} \quad P(A) = \sum_{\substack{i \in I \\ \text{i tq.} \omega_i \in A}} \pi(\omega_i)$$ En particulier : $P(\{\omega_i\}) = \pi(\omega_i)$

TYP On écrit $P(\{\omega\})$, mais il est correct d'écrire $\pi(\omega)$. Incorrect : $\pi(\boxed{\{\omega\}})$

n'est pas un élém. de Ω, mais un sous-ensemble.

■ Définition 14 .... [Probabilité uniforme - équiprobabilité]

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. On dit qu'on a équiprobabilisé Ω.

But : montrer que $P(N) = 0$.

rappel : • l'événement impossible est : $\emptyset$
(pas de $\omega \in \emptyset$).

• N n'est pas l'événement impossible

$N \underset{\text{exple 2}}{=} \{(\text{face}, \text{face}, \text{face}, \dots)\}$ $\#N = 1$

N est un événement élémentaire.

$\boxed{N \neq \emptyset}$

1. Stratégie : . montrer que $N \subset C_n$
. puis prop 1 : $P(N) \leq P(C_n)$
. puis calcul de $P(C_n)$

D'après exemple 2 $$N = \bigcap_{k=1}^{+\infty} F_k$$

$$C_n = \bigcap_{k=1}^{n} F_k$$

D'après Q6 et Q4a :

$$N \subset C_n$$

donc $$\boxed{P(N) \leq P(C_n)}$$

Calculons $P(C_n)$: T7

$$C_n = F_1 \cap F_2 \cap \dots \cap F_n$$

donc

$$P(C_n) = P(F_1 \cap F_2 \cap \dots \cap F_n)$$
$$\cancel{= P(F_1) \cap P(F_2) \cap \dots \cap P(F_n)}$$

par mutuelle indépendance des lancers :

$\cancel{\frac{1}{2} \cap \frac{\pi}{4}}$  $\left\{\frac{1}{2}\right\} \cap \left\{\frac{\pi}{4}\right\}$ 

$$= P(F_1) \times P(F_2) \times \dots \times P(F_n)$$
$$P(C_n) = \frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2}$$
$$= \frac{1}{2^n}$$

Conclusion : $\forall n \geq 1 \quad P(N) \leq \frac{1}{2^n}$ (*)

2. Comme (*) est vraie dès le rang 1, on peut y faire $n \to +\infty$:

$$P(N) \leq 0.$$

Comme $P(N) \geq 0$: $\boxed{P(N) = 0}$

■ Exemple 4.
Un dé fou est un dé possédant une infinité de faces numérotées $0, 1, 2, 3, \ldots$ et truqué : il existe un réel $\lambda > 0$ tel que pour tout entier k, la probabilité d'observer k est à $\lambda 2^{-k}$.

1. Déterminer la valeur de λ.
2. Calculer la probabilité d'observer un score pair en lançant le dé fou.

3 Probabilité conditionnelle

A) Mesure conditionnelle

■ Définition 15 .... [mesure conditionnelle sous l'hypothèse H]
Si $H \in \mathscr{T}$ est tel que $P(H) \neq 0$, la fonction notée P_H définie sur $\mathscr{T}$ par :

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

est une mesure de probabilité sur Ω appelée mesure conditionnelle sous l'hypothèse/observation H.

TYP Fonction La probabilité conditionnelle est une probabilité au sens de `Déf.` 8.

■ Remarque 6.

1. La fonction P_H vérifie donc toutes les propriétés énoncées en **2.A)**. En particulier, $P_H(\bar{A}) = 1 - P_H(A)$.
2. La fonction P_Ω est simplement la fonction P.
3. Par définition, et même si $P(H) = 0$, il est vrai que $P(A \cap H) = P_H(A) \times P(H)$.
4. ☠ $A|H$, ou «A sachant H» n'a **aucun sens** : ni $A|B$, ni «A sachant B» ne sont des évènements : il n'y a aucun sens à exprimer ces notions de *sachant que* en dehors des fonctions P et P_H, car on parle toujours d'un même évènement A, *mais* on *évalue* avec deux fonctions différentes ses chances de réalisation : $P(A)$ ou $P_H(A)$.
5. C'est pour cela que $P(A)$ 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.

B) Formules usuelles

■ Théorème 2 .... [*Formule des probas composées*]

Soit $n \in \mathbf{N}^\star$ et $A_1, \ldots, A_n$ n évènements d'un espace probabiilisé. On pose $\hat{A}_0 = \Omega$ et $\hat{A}_k = \bigcap_{j=1}^{k} A_j$ pour tout $k \in \{1 \ldots n\}$. Alors : $P(\hat{A}_n) = \prod_{k=1}^{n} P(A_k|\hat{A}_{k-1}) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_2 \cap A_1) \times \cdots \times P(A_n|\hat{A}_{n-1})$.

T3 **Quand utiliser la formule des probabilités composées ?**

Formule utilisée pour calculer la probabilité d'une succession chronologique d'évènements, c-à-d. d'un **évènement réalisé par un seul chemin sur l'arbre probabiliste.** La formule est considérablement simplifiée les évènements de cette succession sont mutuellement indépendants.

Exemple 4 : une observation = un entier ≥ 0.

donc $\Omega = \{\text{observations}\} = \mathbb{N}$.

on a d'après l'énoncé : $\pi : \mathbb{N} \longrightarrow \mathbb{R}$
$k \mapsto \frac{\lambda}{2^k}$.

C'est une distribution de masse sur Ω

ssi ① $\forall k \in \mathbb{N} \quad \pi(k) \geq 0$

Ce qui équivaut à $\lambda \geq 0$.

vient de déf 13

② $\sum_{k=0}^{+\infty} \pi(k)$ existe et vaut 1

La série de t.g. $\frac{\lambda}{2^k}$ est géométrique de raison $\frac{1}{2} \in]-1,1[$. Donc elle converge. On calcule (la somme existe) $\sum_{k=0}^{+\infty} \frac{\lambda}{2^k} = \lambda \cdot \frac{1}{1-\frac{1}{2}} = 2\lambda$.

d'où $2\lambda = 1$

$$\boxed{\lambda = \frac{1}{2}}$$

2_ Soit A : "le score observé est pair".

$$A = \{0, 2, 4, 6, 8, 10, \dots\}$$
$$= \{2n \quad n \in \mathbb{N}\}$$

On peut écrire : $A = \bigcup_{k=0}^{+\infty} \{2k\}$

Ces événements (élémentaires) étant 2 à 2 incompatibles,

T_2 par σ-additivité, la série de t.g. $P(\{2k\})$ converge, et sa somme vaut $P(A)$.

$$P(A) = \sum_{k=0}^{+\infty} P(\{2k\}) = \sum_{k=0}^{+\infty} \pi(2k) = \sum_{k=0}^{+\infty} \frac{1}{2} \times \frac{1}{2^{2k}}$$

$$= \frac{1}{2} \sum_{k=0}^{+\infty} \frac{1}{4^k} = \frac{1}{2} \cdot \frac{1}{1-\frac{1}{4}} = \frac{1}{2} \cdot \frac{4}{3}$$

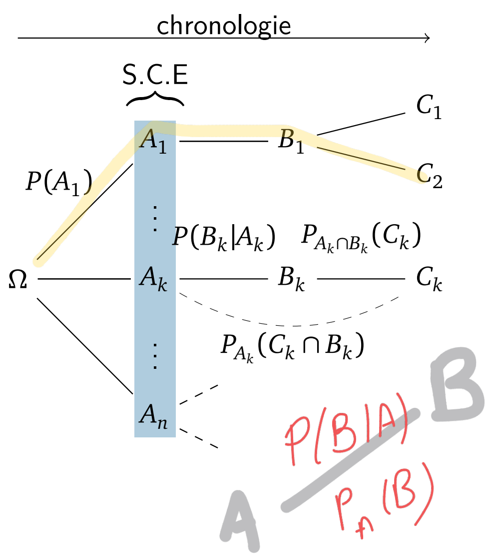
$$\boxed{P(A) = \frac{2}{3}}$$

P(A)

$A \mapsto Q(A) = P(A \cap H) \geq 0$

$Q(\Omega) = P(\Omega \cap H)$
$= P(H)$

T_4 Arbres probabilistes : mode d'emploi

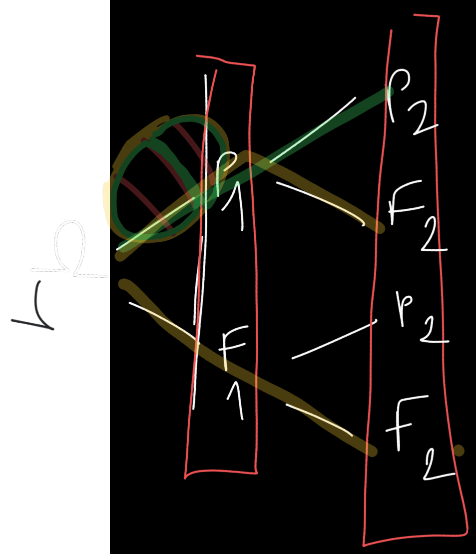


1. **Chemin :** le chemin [Ω—$\bullet$—$\star$—$\circ\ldots$] sur l'arbre représente l'évènement [$\bullet \cap \star \cap \circ \cap \ldots$].
2. **Probabilité conditionnelle :** le poids d'un chemin de racine $\bullet$, mettons : $\bullet - \circ$ est égal à la probabilité conditionnelle $P_\bullet(\circ)$ (En particulier : si $\bullet = \Omega$, on calcule les probabilités *a priori*, voir **Rem.4**).
3. **Formule des probabilités composées :** les poids se multiplient le long des branches (composition des poids).
4. **Systèmes complets :** tout niveau d'arborescence est un S.C.E (ou un S.Q.C.E).
5. **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.
6. ☠ Un évènement ne peut jamais apparaître à deux niveaux d'arborescence distincts. Exemple : suite de pile ou face : $P-P-P$: impossible, mais $P_1-P_2-P_3$: ok.

■ Théorème 3 .... [*Formule des probas totales*]

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

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



T_5 Quand et comment utiliser la formule des probas totales ?

1. Formule adaptée pour calculer la probabilité d'un évènement **se réalisant par plusieurs chemins sur l'arbre.**
2. **a)** La convergence de la série n'est pas à prouver : elle fait partie des conclusions du théorème.

   b) Il est *obligatoire* de citer le sqce utilisé.
3. En pratique, **il est obligatoire** d'écrire quelque chose comme : *D'après la formule de probabilités totales appliquée avec le* sqce $(A_n)_{n\geq 1}$, *la série de t.g.* $P_{A_n}(E)$ *converge, et sa somme vaut* $P(E)$:

$$P(E) = \sum_{n=1}^{\infty} P_{A_n}(E)$$

cf exemple 5

T6 **Quand utiliser la formule de Bayes ?**

Formule permettant de renverser causes et effets dans le calcul stochastique. **S'applique lorsque le chemin $H-A$ ne suit pas la chronologie sur l'arbre.** Le dénominateur $P(A)$ se calcule dans la plupart des cas par **probabilités totales.**

■ Exemple 5.
On dispose d'une infinité d'urnes $U_0, U_1, \ldots$ Chaque urne contient exactement une pièce d'or, et d'autres babioles, si bien que la probabilité de tirer la pièce d'or dans U_k est, pour tout entier k, $\frac{1}{ek!}$. Un lutin joue au jeu suivant :

— Il lance le dé fou. (celui de l'exemple 4)

— Il pioche dans l'urne dont il obtient le numéro par le dé fou.

Calculer la probabilité que le lutin tire une pièce d'or.

■ Théorème 4 .... [*Formule de Bayes*]
Si A et H sont deux évènements de probabilité non nulle :

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

4 Mutuelle indépendance d'évènements ≠ incompatibles

■ Définition 16 .... [Évènements indépendants]
Deux évènements A et B sont $(P-)$indépendants si et seulement si l'une des conditions suivantes équivalentes est vraie :

1. $P(A \cap B) = P(A) \times P(B)$
2. $P(A|B) = P(A)$
3. $P(B|A) = P(B)$

Si on change de fonction P, on peut perdre l'indépendance

■ Remarque 7.
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**.

■ Définition 17 .... [Évènements mutuellement indépendants]
Les évènements $A_1, \ldots A_n$ 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 $i < j < k$: $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$,

$\vdots$

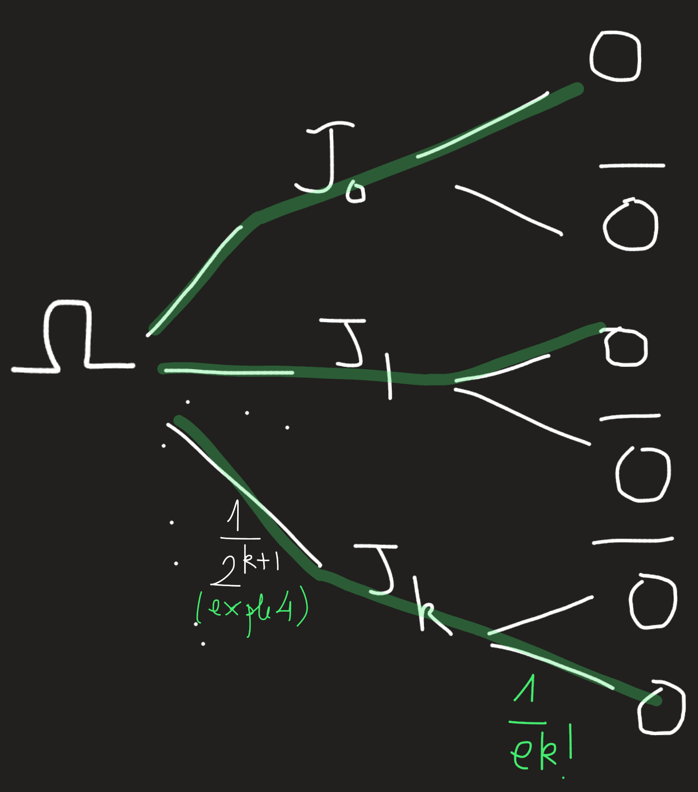
Ck) $\forall i_1, \ldots, i_k$ tels que $i_1 < \cdots < i_k$: $P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \times \cdots \times P(A_{i_k})$,

$\vdots$

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

■ Remarque 8.

1. Dans **Ck**), $\binom{n}{k}$ conditions. Pour n évènements, il y a au total $2^n - n - 1$ conditions dans **Déf. 17** (ex : 26 conditons pour 5 évènements).
2. La mutuelle indépendance est rarement à prouver : elle est soit une hypothèse, soit une conséquence du modèle.



Soit :

J_k : "le lutin joue avec l'urne k"

O : "le lutin tire une pièce d'or"

D'après la formule des probas totales avec le SQCE $(J_k)_{k\in\mathbb{N}}$, la série de t.g $P_{J_k}(O)P(J_k)$ converge et sa somme vaut $P(O)$:

$$P(O) = \sum_{k=0}^{+\infty} P_{J_k}(O)P(J_k)$$

$$= \sum_{k=0}^{+\infty} \frac{1}{ek!} \times \frac{1}{2^{k+1}}$$

$$= \frac{1}{2e} \sum_{k=0}^{+\infty} \frac{1}{2^k k!}$$

$$\underset{x=\frac{1}{2}}{=} \frac{1}{2e} \sum_{k=0}^{+\infty} \frac{x^k}{k!}$$

$$= \frac{1}{2e} e^{1/2} = \frac{\sqrt{e}}{2e} = \frac{1}{2\sqrt{e}}$$

CH 8
Thm 2

■ **Définition 18** .... [indépendance d'une famille infinie 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 au sens de la **Déf. 17**.

■ **Proposition 2** .... [Indépendance et contraire]
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 Comment calculer la probabilité d'un évènement E ?

1. ☠ Le calcul ne commence *jamais* par «$P(E) = \ldots$», car on raisonne *toujours* sur les évènements : on part de E que l'on décompose à l'aide de **T1**.

On commence par $E = \ldots$

2. **Pour les réunions $\bigcup$.**
   a) Réunions finies : additivité finie pour une réunion finie d'évènements deux à deux incompatibles. Sinon, formule du crible.
   b) Réunions infinies : σ-additivité pour une réunion dénombrable d'évènements deux à deux incompatibles.
   c) Penser aussi à la formule des probabilités totales.

   → En passant aux probabilités, les réunions deviennent alors **a)** des sommes finies de nombres, ou **b)** des sommes de séries convergentes.
3. **Pour les intersections $\bigcap$.**
   a) Intersections finies : formule des probabilités composées. Si les évènements de l'intersection sont mutuellement indépendants, l'expression devient moins lourde.
   b) Intersections infinies : *a priori*, on ne peut que **majorer** la probabilité $\pi = P\left(\bigcap_{k=1}^{+\infty} A_k\right)$. On part de $\forall n \geq 1 \quad \bigcap_{k=1}^{+\infty} A_k \subset \bigcap_{k=1}^{n} A_k$, puis en passant aux probabilités : $\forall n \geq 1 \quad \pi \leq P\left(\bigcap_{k=1}^{n} A_k\right)$. On conclut avec le théorème des gendarmes la plupart du temps.

   Cas typique exo 1

   → En passant aux probabilités, les intersections deviennent alors des **produits** de nombres.
4. ☠ Ne pas confondre évènements indépendants et incompatibles !

■ **Exercice 2.**
Soit $p \in]0,1[$. Une pièce truquée ~~donne pile~~ avec probabilité p. On lance cette pièce une infinité de fois. Les lancers sont supposés mutuellement indépedants. On s'intéresse aux longueurs des lancers amenant une même face : on dit que la première série est de longueur $n \in \mathbf{N}^{\star}$ si :

suites de

— Les lancers $1,2,3,\ldots,n$ donnent le même côté de la pièce.

— Le $n+1$-ème lancer donne l'autre côté de la pièce.

On note L_n l'évènement : «la première série est de longeur n». Calculer $P(L_n)$.

Exo2 : Soit $n \geq 1$.

- On note pour tout $k \geq 1$: F_k : "le k^e lancer donne face".

- T1 $L_n = \underbrace{(F_1 \cap \ldots \cap F_n \cap \overline{F_{n+1}})}_{A} \cup \underbrace{(\overline{F_1} \cap \ldots \cap \overline{F_n} \cap F_{n+1})}_{B}$

T7 $\overset{2a)}{=} A \cup B$

- En passant aux probas :

$$P(L_n) = P(A \cup B)$$

Or A et B sont incompatibles car au 1er lancer on a soit pile soit face.

Par additivité finie :

$$P(L_n) = P(A) + P(B).$$

- Calcul de $P(A)$:

T7 3a) Comme les lancers de pièce sont mutuellement indép :

$$P(A) = P(F_1) \times \ldots \times P(F_n) \times P(\overline{F_{n+1}})$$

AN : $P(A) = (1-p)^n \times p$.

de même $P(B) = p^n(1-p)$

$$\boxed{P(L_n) = p^n(1-p) + (1-p)^n p}$$

5 Notion de variable aléatoire sur un espace probabilisé

A) Définition

■ **Définition 19** .... [Variable aléatoire]

Soit $(\Omega, \mathscr{T})$, un espace probabilisable. On appelle variable aléatoire sur Ω toute *fonction* X définie sur Ω telle que pour tout intervalle I de $\mathbf{R}$, l'*ensemble* noté

$$[X \in I] = \{\omega \in \Omega \quad X(\omega) \in I\}$$

vérifie : $[X \in I] \in \mathscr{T}$, autrement dit $[X \in I]$ est un évènement.

■ **Remarque 9.**

1. Intérêt de cette notion : les valeurs prises par X s'interprètent comme des mesures effectuées sur les observations. La définition garantit que les observations des valeurs prises par X sont bien des évènements. Elles peuvent donc être analysées en termes de probabilités.

2. $[X \in I]$ **est une notation pour désigner un** **TYP ensemble** . $[X \in I]$, **n'est pas une** **TYP assertion** **exprimant une appartenance,** $[X = a]$ **n'est pas une égalité,** $[X > a]$ **n'est pas une inéquation, ni une inégalité.**

■ **Exercice 3.**
Traduire en langage probabiliste : **1.** $X = 2$, **2.** $[X = 2]$, **3.** $X \in [2, 5]$.

B) Indépendance de variables aléatoires

■ **Définition 20** .... [Variables aléatoires indépendantes]

1. Soit $n \geq 2$ un entier et $X_1, \ldots X_n$ n variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{T}, P)$. On dit que $X_1, \ldots X_n$ sont indépendantes si pour tous intervalles $I_1, \ldots, I_n$ de $\mathbf{R}$:
les évènements $[X_1 \in I_1], \ldots [X_n \in I_n]$ sont indépendants.

2. Soit $(X_k)_{k \geq 1}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{T}, P)$. On dit que la suite de variables (X_k) est une suite de variables indépendantes si pour tout entier $n \geq 2$, et tout sous ensemble fini $\{i_1, \ldots, i_n\}$ de n entiers strictement positifs, les variables $X_{i_1}, \ldots X_{i_n}$ sont indépendantes

C) Fonction de répartition d'une variable aléatoire réelle

■ **Définition 21** .... [Fonction de répartition]
Soit X une variable aléatoire sur $(\Omega, \mathscr{T}, P)$. La fonction de répartition de X est la fonction notée F_X définie sur $\mathbf{R}$ par :

$$\forall t \in \mathbf{R} \quad F_X(t) = P(X \leq t)$$

■ **Proposition 3** .... [Propriétés universelles des fonctions de répartition]
Une fonction de répartition vérifie les propriétés suivantes :

1. Elle est définie sur $\mathbf{R}$.
2. Elle est positive sur $\mathbf{R}$.
3. Elle est croissante sur $\mathbf{R}$.
4. Elle admet des limites en $-\infty$ et en $+\infty$ égales respectivement à 0 et 1.

jamais besoin de le prouver si on sait que F est une fonc. de rép.

Exercice 3

1. $X=2$

pas synonymes. N'enlevez pas les [ ].

2. $[X=2]$

optionnel car cela surcharge

3. $[X\in[2,5]]$

$X=2$ La fonction X est cte égale à 2 : $X \sim \mathcal{C}(2)$

$[X=2]$ À l'issue de la mesure faite sur l'observation, la mesure obtenue est 2.

$X\in[2,5]$ la mesure faite est comprise entre 2 et 5

On lance une pièce : X = VAR qui vaut 0 si on obtient face 1 sinon.

"On a obtenu pile donc $X=1$" ✗

"on a obtenu pile donc $\begin{cases} X(\omega)=1 \\ [X=1] \text{ est réalisé} \end{cases}$" ✓.

$$
\begin{tikzcd}
\Omega \arrow[r, bend left] & \omega \in \Omega \arrow[d] \\
& T(\omega) \in \mathbb{R}
\end{tikzcd}
$$