

CH2 – Théorie basique des probabilités

Plan du chapitre

1	Vocabulaire des probabilités	3
	A) Ce qu'est une expérience aléatoire	3
	B) Dictionnaire biligne probas/théorie des ensembles	4
2	Mesure de probabilité	6
	A) Propriétés axiomatiques d'une probabilité	6
	B) Construction de probabilités	7
3	Probabilité conditionnelle	8
	A) Mesure conditionnelle	8
	B) Formules usuelles	8
4	Mutuelle indépendance d'évènements	10
5	Notion de variable aléatoire sur un espace probabilisé	12
	A) Définition	12
	B) Indépendance de variables aléatoires	12
	C) Fonction de répartition d'une variable aléatoire réelle	12

Liste des définitions

Déf.1	Ensemble infini et dénombrable	3
Déf.2	Ensemble au plus dénombrable	3
Déf.3	Expérience	3
Déf.4	Expérience aléatoire	3
Déf.5	Univers associé à une expérience	3
Déf.6	Observation - Évènement élémentaire	3
Déf.7	évènement, certain, impossible, implique, contraire, incompatibles, SCE	4
Déf.8	Tribu (d'évènements)	6
Déf.9	Mesure de probabilité ou probabilité	6
Déf.10	Espace probabilisable - espace probabilisé	6
Déf.11	SQCE, quasi-certain/presque sûr, quasi-impossible	7
Déf.12	Distribution/fonction de masse sur Ω	7
Déf.13	Probabilité uniforme - équiprobabilité	7
Déf.14	mesure conditionnelle sous l'hypothèse H	8
Déf.15	Évènements indépendants	10
Déf.16	Évènements mutuellement indépendants	10
Déf.17	indépendance d'une famille infinie d'évènements	11
Déf.18	Variable aléatoire	12
Déf.19	Variables aléatoires indépendantes	12
Déf.20	Fonction de répartition	12

Faire attention à ce qui suit

Dénombrabilité

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

■ **Remarque 1.**

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

■ **Exemple 1.**

Les ensembles $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{Z}$, $\mathbb{Q}$ 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.

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 réunion/intersection/somme¹ 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** [Expérience]
Processus reproductible aboutissant à un résultat observable ω .

■ **Définition 4** [Expérience aléatoire]
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 .

■ **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.

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.

B) Dictionnaire biligüe probas/théorie des ensembles

■ **Définition 7** .. [évènement, certain, impossible, implique, contraire, incompatibles, SCE]

Notation	Terme ensembliste	Traduction probabi- liste
$\mathcal{T}$	Tribu ($\mathcal{T} = \mathcal{P}(\Omega)$ en sup.)	Ensemble des évènements
Ω	partie pleine	univers
$\emptyset$	ensemble vide	évènement impossible
$\omega \in \Omega$	élément de Ω	ω est une observation
$A \in \mathcal{T}$	élément de $\mathcal{T}$	A est un évènement
$\omega \in A$	ω appartient à A	ω est une réalisation de l'évènement A
$A \subset B$	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]
$A \cap B = \emptyset$	A et B sont disjoints	A et B sont incompatibles
$(A_i)_{i \in I}$ t.q. $\bigcup_{i \in I} A_i = \Omega$ $i \neq j \Leftrightarrow A_i \cap A_j = \emptyset$	$(A_i)_{i \in I}$ est une partition de Ω	$(A_i)_{i \in I}$ est un système complet d'évènements (SCE)

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 faudra donc toujours commencer, avant de calculer des probabilités, 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.

1. Lister les observations possibles.
2. En déduire l'univers Ω .
3. Expliciter les événements 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. »

■ 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. Pour tout entier $k \geq 1$, on note F_k l'évènement : le k -ème lancer donne face. Est-ce que F_1 est un événement élémentaire ?
3. Exprimer à l'aide des événements F_k les événements 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. »

2 Mesure de probabilité

■ Définition 8 [Tribu (d'évènements)]

Les tribus sont les structures garantissant que les opérations standard sur les évènements donnent encore des évènements, à savoir les opérations ^{c} , $\bigcup_{i \in I}$, et $\bigcap_{i \in I}$. Autrement dit, ce sont des ensembles d'évènements stables par les opérations précédentes.

A) Propriétés axiomatiques d'une probabilité

■ Définition 9 [Mesure de probabilité ou probabilité]

Soit Ω un univers, équipé d'une tribu $\mathcal{T}$. Une *probabilité* sur Ω est une fonction P définie sur $\mathcal{T}$ vérifiant les conditions suivantes :

1. *Normalisation.* : $P(\Omega) = 1$.
2. *Positivité.* $\forall A \in \mathcal{T} \quad P(A) \geq 0$.
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 $P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$.

■ Remarque 4.

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

■ Remarque 5.

Attention au sens du mot *probabilité* :

1. **TYP Fonction** Probabilité = La fonction définie sur $\mathcal{T}$ (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 $A \in \mathcal{T}$ 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 incompatibles.
2. On ajoute ensuite : « par σ -additivité la série de t.g. $P(A_n)$ est convergente et sa somme vaut $P\left(\bigcup_{n=1}^{+\infty} A_n\right)$: »

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

■ Définition 10 [Espace probabilisable - espace probabilisé]

Un univers Ω équipé d'une tribu d'évènements $\mathcal{T}$ est dit probabilisable. Un espace probabilisable équipé d'une fonction de probabilité P définie sur $\mathcal{T}$ fait du triplet $(\Omega, \mathcal{T}, P)$ un espace probabilisé.

■ Proposition 1 [Propriétés d'une probabilité P]

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

■ Définition 11 [SQCE, quasi-certain/presque sûr, quasi-impossible]

Notation	Traduction ensembliste	Traduction probabiliste
$\bullet P\left(\bigcup_{i \in I} A_i\right) = 1$ $\bullet i \neq j \Leftrightarrow A_i \cap A_j = \emptyset$	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)
$P(A) = 1$		A est un évènement quasi-certain/presque sûr
$P(A) = 0$		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 $n \geq 1$: $P(N) \leq \frac{1}{2^n}$. Conclure.

B) Construction de probabilités

■ Définition 12 [Distribution/fonction de masse sur Ω]

Soit $\Omega = \{\omega_i \mid i \in I\}$ 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 : $\sum_{i \in I} p_i = 1$.

■ 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 $(\Omega, \mathcal{T})$ par :

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

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

■ Définition 13 [Probabilité uniforme - équiprobabilité]

Si l'ensemble I (et donc Ω) est fini, en prenant π constante égale à $\frac{1}{\#\Omega}$, on obtient comme mesure P la mesure appelée probabilité uniforme ou équiprobabilité sur Ω . On dit qu'on a équiprobabilisé Ω .

■ Exemple 4.

Un dé fou est un dé possédant une infinité de faces numérotées $0, 1, 2, 3, \dots$ 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 14** [mesure conditionnelle sous l'hypothèse H]

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

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

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 événements : 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 événement 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 $n \in \mathbb{N}^*$ et $A_1, \dots, A_n$ n événements d'un espace probabilisé. On pose $\hat{A}_0 = \Omega$ et $\hat{A}_k = \bigcap_{j=1}^k A_j$ pour tout $k \in \{1 \dots n\}$. Alors : $P(\hat{A}_n) = \prod_{k=1}^n P(A_k | \hat{A}_{k-1}) = P(A_1) \times P(A_2 | A_1) \times P(A_3 | A_2 \cap A_1) \times \dots \times P(A_n | \hat{A}_{n-1})$.

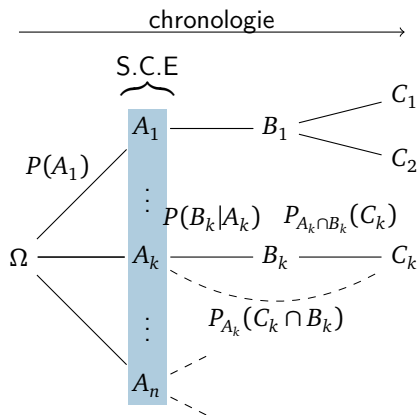
T₃

Quand utiliser la formule des probabilités composées ?

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

T₄

Arbres probabilistes : mode d'emploi



- 1. Chemin :** le chemin $[\Omega \rightarrow \bullet \rightarrow \star \rightarrow \circ \dots]$ sur l'arbre représente l'évènement $[\bullet \cap \star \cap \circ \dots]$.
- 2. Probabilité conditionnelle :** le poids d'un chemin de racine $\bullet$, mettons : $\bullet \rightarrow \circ$ 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'arborescence, $P(\bullet)$ 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

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

T₅

Quand et comment utiliser la formule des probas totales ?

- Formule adaptée pour calculer la probabilité d'un évènement **se réalisant par plusieurs chemins sur l'arbre**.
- La convergence de la série n'est pas à prouver : elle fait partie des conclusions du théorème.
 - Il est *obligatoire* de citer le sqce utilisé.
- En pratique, **il est obligatoire** d'écrire quelque chose comme : *D'après la formule de probabilités totales appliquée avec le sqce $(A_k)_{k \geq 1}$, [la série de t.g. $P_{A_k}(E)P(A_k)$ converge, et sa somme vaut $P(E)$]* :

$$P(E) = \sum_{k=1}^{\infty} P_{A_k}(E)P(A_k)$$

- Bien évidemment, si le sqce est fini, la somme apparaissant ne contient qu'un nombre fini de termes, il n'y a donc pas de série dans ce cas, donc on n'écrit pas la partie [...] de **3.a**), et on passe directement à :

$$P(E) = \sum_{k=1}^n P_{A_k}(E)P(A_k).$$

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**, et a dû être calculé dans une question antérieure.

■ Exemple 5.

On dispose d'une infinité d'urnes $U_0, U_1, \dots$. 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.
- 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 événements de probabilité non nulle :

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

4 Mutuelle indépendance d'évènements**■ Définition 15** [Évènements indépendants]

Deux événements 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. $P(B|A) = P(B)$

■ 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 16 [Évènements mutuellement indépendants]

Les événements $A_1, \dots, A_n$ sont mutuellement indépendants si et seulement si :

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

C3) $\forall i, j, k$ tels que $i < j < k$: $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$,

⋮

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

⋮

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

■ Remarque 8.

1. Dans Ck), $\binom{n}{k}$ conditions. Pour n événements, il y a au total $2^n - n - 1$ conditions dans **Déf. 17** (ex : 26 conditions pour 5 événements).

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

■ **Définition 17** [indépendance d'une famille infinie d'évènements]

La famille $\mathcal{F}$ est une famille d'évènements mutuellement indépendants si toute sous-famille finie d'évènements extraite de $\mathcal{F}$ 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 *jamaïs* par « $P(E) = \dots$ », car on raisonne *toujours* sur les évènements : on part de E que l'on décompose à l'aide de **T1**.

2. **Pour les réunions $\cup$.**

- 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 $\cup$ deviennent alors a) des sommes (Σ) finies de nombres, ou b) des sommes de séries convergentes ($\sum_{k=1}^{\infty}$).

3. **Pour les intersections $\cap$.**

- 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.

→ En passant aux probabilités, les $\cap$ d'évènements deviennent alors des **produits** (Π) 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 indépendants. On s'intéresse aux longueurs des lancers amenant une même face : on dit que la première série est de longueur $n \in \mathbb{N}^*$ si :

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

■ 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.