

PROBABILITÉS



Plan



- Chapitre 1: Introduction aux probabilités
- Chapitre 2 : Variables aléatoires discrètes
- Chapitre 3 : Variables aléatoires continues

1.0 Introduction



FIGURE 1 – A. Kolmogorov

- Les probabilités sont l'étude du hasard et de l'incertain. Elle permettent de donner un cadre formel et rigoureux aux nombreux phénomènes physique aléatoires.
- Comme le dit Emile Borel, un des fondateurs de la théorie de la mesure, le hasard est une science; elle provient à l'origine de l'étude des jeux de hasard, notamment des jeux de dés.
- Bien que plusieurs grands mathématiciens, dont Pascal et Laplace, ont tenté de formaliser ces phénomènes, il faut attendre le XXème siècle pour que la théorie moderne des probabilités, telle que nous la connaissons aujourd'hui, soit élaborée:

En effet, c'est grâce aux travaux sur la théorie de l'intégration de Borel et Lebesgue, plus générale que celle de Riemann, que les prémisses de cette théorie furent posés. Par la suite ce sera grâce à Lévy, passant par Itô, Kolmogorov qu'elle verra le jour.

1.1 L'Univers

- **Définition** : On appelle expérience aléatoire une expérience renouvelable et qui, renouvelée dans des conditions identiques, ne donne pas forcément le même résultat.

Exemples: lancer un dé, lancer une pièce de monnaie, tirer une boule d'une urne, ...

- **Définition** : L'ensemble des issues possibles d'une expérience aléatoire donnée se nomme l'univers (ou ensembles des issues); On le note Ω . Un élément de Ω est donc une issue, et on la représente par ω .

Exemples dans l'expérience aléatoire où on :

- lance un dé, $\Omega = \{1, 2, 3, 4, 5, 6\}$;
- lance une pièce de monnaie, $\Omega = \{P, F\}$;

1.2 Algèbre des événements

- **Définition** : Un événement aléatoire est une sous-expérience d'une expérience aléatoire donnée. En théorie des ensembles un événement est un sous-ensemble de Ω .

Exemple : dans l'expérience aléatoire où on lance un dés, $\Omega = \{1, 2, 3, 4, 5, 6\}$; parmi les événements issues de cette expérience, on trouve A « obtenir un nombre pair »; explicitement $A = \{2, 4, 6\}$.

- **Définition** : Une famille \mathcal{A} de parties de l'univers Ω est une tribu, si elle satisfait les trois propriétés suivantes :

- $\Omega \in \mathcal{A}$
- Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$
- Soit $(A)_{i \in I}$, une famille dénombrable¹ d'éléments de \mathcal{A} , alors $\bigcup_{i \in I} A_i \in \mathcal{A}$

- **Exemple** : $\mathcal{P}(\Omega)$, $\{\Omega, \emptyset\}$ et $\{A, A^c, \Omega, \emptyset\}$ sont des tribus où A est un événement.

1.2 Algèbre des événements

Propriétés 1.3.1. Soit \mathcal{A} une tribu d'un univers Ω . Les propriétés suivantes sont des conséquences directes de la définition :

1. $\emptyset \in \mathcal{A}$.
2. Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de \mathcal{A} alors $\bigcap_{n=1}^{+\infty} A_n \in \mathcal{A}$.
3. Si $(A_i)_{0 \leq i \leq N}$ est une suite finie de N éléments de \mathcal{A} alors $\bigcup_{i=0}^N A_i \in \mathcal{A}$.
4. Si $(A_i)_{0 \leq i \leq N}$ est une suite finie de N éléments de \mathcal{A} alors $\bigcap_{i=0}^N A_i \in \mathcal{A}$.

- Puisque $\Omega \in \mathcal{T}$, alors $\phi \in \mathcal{T}$
- Il suffit de prendre $\forall n \in \mathbb{N} \quad B_n = A_n^c$ et ce pour tout $A_n \in \mathcal{T}$
- Il suffit de concevoir une suite d'ensemble statique à partir d'un certain rang, par exemple $\forall m \geq n, A_m = A_n$ et ce pour retrouver le cas dénombrable
- On procède de la même manière que 2.

1.3 La notion de probabilité

Probabilité expérimentale = inductive = fréquentiste = a posteriori

La probabilité expérimentale ou inductive ou fréquentiste ou a posteriori est déduite de toute la population concernée.

Exemple: si sur une population d'un million de naissances, on constate 530000 garçons et 470000 filles, on dit que $P[\text{garçon}] = 0.53$.

La vision fréquentiste repose sur la *loi des grands nombres*, établie pour la première fois par J. Bernoulli en 1713, fournit une définition 'pratique' de la notion de probabilité.

Cette loi stipule que si on répète un grand nombre n de fois une épreuve, la fréquence f avec laquelle on observe la survenue d'un événement tend (quand $n \rightarrow \text{infini}$) vers une limite qui est définie comme probabilité de cet événement.

Exemple: en lançant un nombre n de fois une pièce 'parfaite' de monnaie, la fréquence f de 'pile' tendra vers $1/2$ au fur et à mesure que n augmentera.

N. lancés (n)	N. piles en n lancés	Fréquence
10	6	0.60000
100	52	0.52000
1000	463	0.46300
10000	4908	0.49080
100000	49573	0.49573

1.3 La notion de probabilité

Probabilité théorique = déductive = a priori



La probabilité théorique ou déductive ou a priori est connue grâce à l'étude du phénomène sous-jacent sans expérimentation. Il s'agit donc d'une connaissance *a priori* par opposition à la définition précédente qui faisait plutôt référence à une notion de probabilité *a posteriori*.

Exemple 1: dans le cas classique du dé parfait, on peut dire, sans avoir à jeter un dé, que $P[\text{"obtenir un 4"}] = 1/6$.

Exemple 2: dans le cas classique d'une pièce de monnaie non truquée, on peut dire, sans avoir la lancée, que $P[\text{"obtenir un pile"}] = 1/2$.

1.3 La notion de probabilité

de la *a posteriori* à la probabilité *a priori* / problème



Comme il n'est pas toujours possible de déterminer des probabilités *a priori*, on est souvent amené à réaliser des expériences. Il faut donc pouvoir passer de la première à la deuxième solution. Ce passage est supposé possible en terme de limite (*i.e.* avec une population dont la taille tend vers la taille de la population réelle).

Remarque: dans la conception fréquentiste, il est impossible de donner une valeur et même au sens à la probabilité d'un événement non répétable du genre « neigera-t-il le 25 octobre 2990 » , ce qui limite le champ d'application du calcul des probabilités.

Il est clair donc qu'il faut être capable de définir la probabilité autrement que par une approche fréquentielle si on veut être capable de parler de probabilité d'événements qui ne peuvent se produire qu'une fois et pour lesquels la répétition de l'épreuve n'a aucun sens (comme c'est le cas dans l'exemple précédent).

1.4 Axiomatique de Kolmogorov

Définition: On appelle espace probabilisable, le couple $(\Omega; \mathcal{T})$, où \mathcal{T} est une tribu de Ω .

Définition: Une probabilité (ou mesure de probabilité 2) sur $(\Omega; \mathcal{T})$, est une application p de \mathcal{T} dans $[0,1]$ vérifiant les trois axiomes de Kolmogorov suivants:

Axiome 1: $\forall A \in \mathcal{T}, 0 \leq p(A) \leq 1$

Axiome 2: $p(\Omega) = 1$

Axiome 3: Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de \mathcal{T} deux à deux incompatibles.

$$\text{On a : } p\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} p(A_n)$$

Définition: On appelle espace probabilisé, le triplé $(\Omega; \mathcal{T}; p)$, où \mathcal{T} est une tribu de Ω et p une probabilité.

1.4 Axiomatique de Kolmogorov

Propriétés : soient A et B deux événements d'un univers Ω .

1. $p(A^c) = 1 - p(A)$

2. $p(\emptyset) = 0$

3. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

4. $A \subseteq B \Rightarrow (p(A) \leq p(B))$ et $(p(B \setminus A) = p(B) - p(A))$

(ce résultat peut être généralisé à plus de 2 événements)

Démonstrations:

1. On a $A \cup A^c = \Omega$, alors $p(A^c) = 1 - p(A)$;

2. *Trivial*

3. On a $A \cup B = (A \setminus A \cap B) \cup B$

alors $p(A \cup B) = p(A \setminus A \cap B) + p(B) = p(A) - p(A \cap B) + p(B)$

4. Puisque $A \subseteq B$, il existe un événement D tel que $A \cap D = \emptyset$ et $A \cup D = B$

Alors $p(A) + p(D) = p(B)$; mais $p(D) \geq 0$, alors $p(A) \leq p(B)$.

1.5 Loi de probabilité uniforme discrète



Afin de formaliser la notion de probabilité, cette sous partie traite l'exemple de la probabilité uniforme discrète.

Vocabulaire: Un ensemble est dit discret , s'il peut être mis en bijection avec une sous-partie de \mathbb{N} . Discret est synonyme de dénombrable.

Exemples:

- Toute ensemble fini est discret;
- \mathbb{N} est discret;
- \mathbb{R} n'est pas discret.

1.5 Loi de probabilité uniforme discrète

Définition : Soit un univers discret fini. La loi de probabilité uniforme discrète, est une probabilité qui associe à chaque élément w de l'univers Ω la même valeur.

Exemple : Prenons l'exemple où on lance un dé équilibré.

L'univers, qui est l'ensemble des issues possibles de cette expérience est donc égal $\Omega = \{1; 2; 3; 4; 5; 6\}$. Soit p la loi de probabilité uniforme discrète. On a :
 $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6$

Exemple : Prenons l'exemple d'un jeu de pile ou face avec une pièce parfaite.

L'univers, qui est l'ensemble des issues possibles de cette expérience est donc égal à $\Omega = \{\text{pile}, \text{face}\}$.

Soit p la loi de probabilité discrète uniforme. On a donc : $p(\text{pile}) = p(\text{face}) = 1/2$

1.5 Loi de probabilité uniforme discrète

Cas général : Soit un univers discret fini d'une expérience et soit la loi de probabilité uniforme discrète.

Si le cardinal de Ω , $|\Omega|$, vaut n alors on a : pour tout événement A , on a
 $p(A) = |A|/n$

Exemple : Prenons l'exemple d'un lancé de deux dés non pipés.

L'univers, qui est l'ensemble des issues de cette expérience, est donc égal à :

$$\Omega = \{1;2;3;4;5;6\} \times \{1;2;3;4;5;6\} \text{ donc } |\Omega| = 36$$

En utilisant la probabilité uniforme discrète, calculons l'événement A : "Obtenir un 7"
; pour obtenir un 7, il faut que la somme des deux dés soit égale à 7. Les issues de vérifiant ceci sont : (1; 6); (2; 5); (3; 4); (4; 3); (5; 2) et (6; 1) donc $|A| = 6$

On en déduit que la probabilité d'obtenir un 7 en lançant deux dés est égale à :

$$p(A) = 6/36 = 1/6$$

1.6 Indépendance

Définition : Soit (Ω, \mathcal{T}, p) un espace probabilisé et soit A et B deux évènements définis sur cet espace.

On dit que A et B sont indépendants si et seulement si $p(A \cap B) = p(A)p(B)$

Exemple : Prenons l'exemple d'un lancé successif de deux dés non pipés.

Notons A l'évènement "obtenir un 5 avec le premier dé" et B l'évènement "obtenir un 3 avec le deuxième dé".

- L'univers des possibilités est $\Omega = \{1; 2; 3; 4; 5; 6\} \times \{1; 2; 3; 4; 5; 6\}$ donc $|\Omega| = 36$
- $A = \{(5; 1); (5; 2); (5; 3); (5; 4); (5; 5); (5; 6)\}$ donc $|A| = 6$ d'où $p(A) = 1/6$
 $B = \{(1; 3); (2; 3); (3; 3); (4; 3); (5; 3); (6; 3)\}$ donc $|B| = 6$ d'où $p(B) = 1/6$
- $A \cap B = \{(5; 3)\}$ donc $p(A \cap B) = 1/36$
- Finalement, A et B sont indépendants

1.6 Indépendance

Proposition: Soit (Ω, \mathcal{T}, P) un espace probabilisé et soit A et B deux évènements indépendants. Alors :

- 1- A et B^c sont indépendants
- 2- A^c et B sont indépendants
- 3- A^c et B^c sont indépendants

Démonstration:

1- on a $p(A \cap B^c) = p(A \setminus A \cap B) = p(A) - p(A \cap B)$

Si A et B sont indépendants, alors

$$p(A \cap B^c) = p(A)(1 - p(B)) = p(A)p(B^c)$$

1.7 Probabilité conditionnelle-composée

- Soient deux évènements A et B réalisés, respectivement, n et m fois au cours de N épreuves. On a donc $p(A)=n/N$ et $p(B)=m/N$. Si de plus A et B sont réalisés simultanément k fois, on a $p(A \text{ et } B)=k/N$.
- la probabilité de l'évènement B sachant que l'évènement A est réalisé est appelée **probabilité conditionnelle de B sachant A** et se note $p(B/A)$. Dans notre cas, on a $p(B/A)=k/n = p(A \text{ et } B)/p(A)$.
- **Définition** : $p(B/A)=p(B \text{ et } A)/p(A)$ et $p(A/B)=p(A \text{ et } B)/p(B)$.
- **Remarque** : la probabilité conditionnelle est nommée aussi la probabilité a posteriori car on sait que l'expérience a été réalisée et un événement a été réalisé; là on fait la mise à jour de la probabilité des autres événements. Dans ce cadre, l'évènement réalisé joue le rôle de l'univers de possibilités.

1.7 Probabilité conditionnelle-composée

□ Conséquences :

Deux évènements A et B sont dits indépendants si $p(A \text{ et } B) = p(A) \cdot p(B)$ ou encore si $p(B/A) = p(B)$ (l'information sur la réalisation de A n'apporte rien à l'évènement B) et $p(A/B) = p(A)$.

- **Remarque :** Il est très important de remarquer que d'écrire $P(A/B)$ n'est qu'une notation. La probabilité conditionnelle doit être vu comme une probabilité prenant en argument l'évènement A. Ainsi l'argument ne dépend pas de B. En particulier, on a :

1- Si A et B sont deux évènements, alors $p(A^c / B) = 1 - p(A / B)$;

2- Si $(A_i)_{1 \leq i \leq n}$ est une famille d'évènements deux à deux incompatibles

tel que $\bigcup_{i=1}^n A_i = \Omega$, alors $p(A) = \sum_{i=1}^n p(A/A_i)p(A_i)$.

1.7 Probabilité conditionnelle-composée : exemple

Exemple: Le personnel d'une entreprise est composé de 80% de femmes, 10% des femmes fument et 30% des hommes fument. Quelle est la proportion p des fumeurs dans cette entreprise ?

Pour répondre à cette question, on aura besoin des événements suivants:

F «femmes » ; H «hommes » ; A «fumeurs»

Alors $p(A)=p(A/F)p(F)+p(A/H)p(H)=0.1 \times 0.8 + 0.3 \times 0.2 = 0.14$

Donc $p=14\%$

1.8 Théorème de Bayes - Probabilités des causes

Théorème 2.2.4 (Théorème de Bayes). Soit $(A_i)_{i \in I}$ une famille d'évènements dénombrable incompatible deux à deux, telle que $\forall i \in I, \mathbb{P}(A_i) > 0$ et :

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = 1$$

On a : $\forall A \in \mathcal{A}$ tel que $\mathbb{P}(A) > 0$, alors $\forall i \in I$:

$$\mathbb{P}(A_i|A) = \frac{\mathbb{P}(A|A_i)\mathbb{P}(A_i)}{\sum_{j \in I} \mathbb{P}(A|A_j)\mathbb{P}(A_j)}$$

Démonstration 2.2.5. En utilisant la définition de la probabilité conditionnelle on a :

$$\begin{aligned} \mathbb{P}(A_i|A) &= \frac{\mathbb{P}(A_i \cap A)}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(A|A_i)\mathbb{P}(A_i)}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(A|A_i)\mathbb{P}(A_i)}{\sum_{j \in I} \mathbb{P}(A|A_j)\mathbb{P}(A_j)} \quad (\text{application de la formule des probabilités totales}) \end{aligned}$$

1.8 Théorème de Bayes - Probabilités des causes : exemple

Exemple : Dans une usine, deux machines M_1 et M_2 produisent des circuits intégrés; dans ce cadre, M_1 produit 70% des circuits. En plus, M_1 produit 20% des pièces défectueuses et M_2 en produit 10%. Quelle est la probabilité pour qu'un objet défectueux ait été fabriqué par la machine M_1 ? On cherche donc $p(M_1 / D)$

Pour répondre à cette question, on aura besoin des événement suivants:

M_1 « la pièce est fabriquée par M_1 »

M_2 « la pièce est fabriquée par M_2 »

D « la pièce est défectueuse »

En appliquant la formule de Bayes, on trouve:

$$\begin{aligned} p(M_1 / D) &= p(D / M_1) \cdot p(M_1) / (p(D / M_1) \cdot p(M_1) + p(D / M_2) \cdot p(M_2)) \\ &= 0.2 \times 0.7 / (0.2 \times 0.7 + 0.3) = \sim 0.82 \end{aligned}$$