2. Notions de probabilités

Définitions historiques

Comment définir la probabilité $P(A)$ d'un événement A à partir d'une expérience aléatoire ?

- *Définition classique* (Laplace, 1819)

$\#A$	nombre de cas équiprobables réalisant A
$\#\Omega$	nombre de cas équiprobables possibles

$$\Longrightarrow P(A) = \frac{\#A}{\#\Omega}$$

- Problèmes :
  - les cas doivent être équiprobables
  - le nombre de cas possible doit être fini
  - la définition dépend du terme "équiprobable"

Exemple : obtenir 7 avec deux dés ➡

Si on s'intéresse à la somme des points obtenus, on a

$$\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \quad \Longrightarrow \quad \#\Omega = 11$$

L'événement $A \equiv$"la somme vaut 7" représente un cas parmi les 11.
On en déduirait donc que

$$P(A) = \frac{\#A}{\#\Omega} = \frac{1}{11}$$

Ce résultat est bien entendu incorrect.

Question : quelle est la valeur correcte de $P(A)$?

Définitions historiques

- *Définition en fréquence*

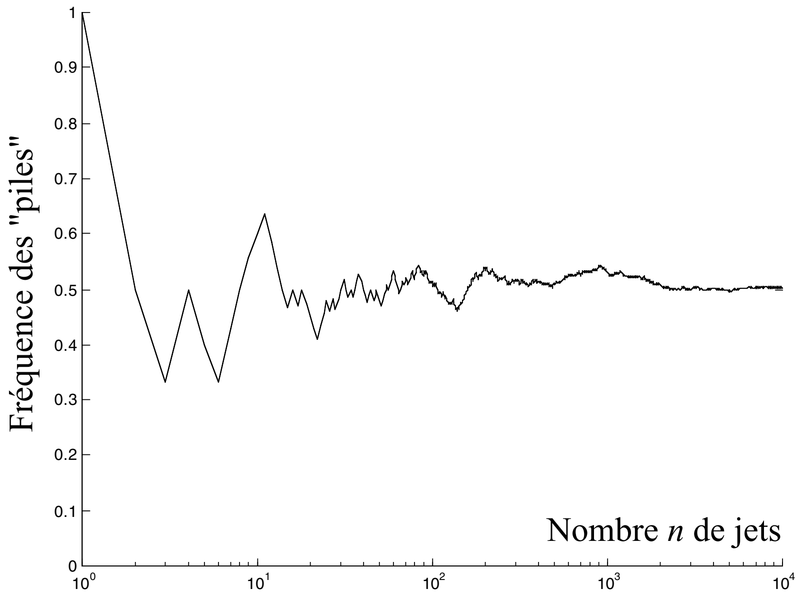
n	nombre de répétitions de l'expérience
$n(A)$	nombre de réalisations de A parmi les n répétitions

$$\Longrightarrow P(A) = \lim_{n\to\infty} \frac{n(A)}{n}$$

- Problèmes : – probabilité définie comme une limite
  – définition peu commode à utiliser

Exemple : fréquence relative des “piles” pour une pièce

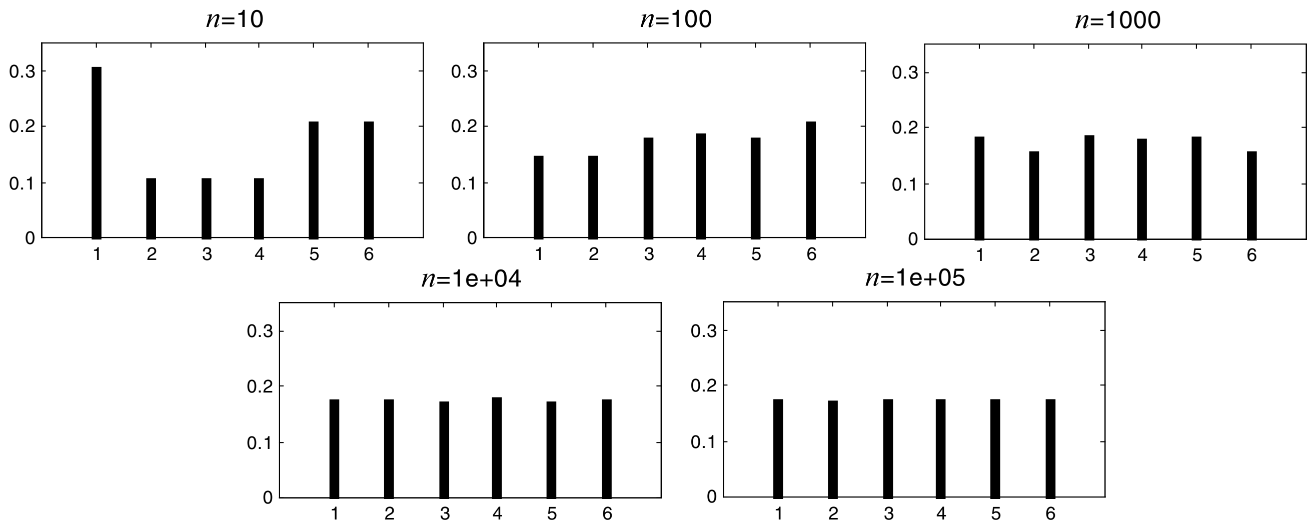
On peut tracer l'évolution de la fréquence relative en fonction de n :



La fréquence relative tend vers la valeur théorique de 1/2.

Exemple : Fréquences relatives des résultats pour un dé

On peut tracer l'évolution de la fréquence relative de chacun des 6 résultats en fonction de n :



Chaque fréquence relative tend vers la valeur théorique de 1/6.

Définition axiomatique (Kolmogorov, 1933)

Axiome = proposition que l'on choisit de considérer comme vraie, sans démonstration, et qui sert de point de départ pour les développements théoriques.

- Axiome 1 : la probabilité est un nombre non négatif
  $P(A) \geq 0$
- Axiome 2 : la probabilité de l'événement certain vaut 1
  $P(\Omega) = 1$
- Axiome 3 : pour deux événements incompatibles, la probabilité de leur union vaut la somme de leur probabilité
  $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$

Définitions

La définition axiomatique permet de dériver aisément toutes les formules nécessaires **uniquement** sur base des 3 axiomes.

Par exemple :

- ☐ $P() = 0$
- ☐ $P(A) \leq 1$
- ☐ $P(A^*) = 1 - P(A)$
- ☐ $P(A) \leq P(B)$ si $A \subseteq B$
- ☐ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Exemple

Messieurs A et B sont des tireurs à l'arc.

- la probabilité que A touche la cible vaut 0.7
- la probabilité que B touche la cible vaut 0.5
- si A et B tirent en même temps, la probabilité qu'ils touchent tous les deux la cible vaut 0.35

Quelle est la probabilité qu'au moins une flèche touche la cible si A et B tirent en même temps ?

Définitions

En utilisant les lois de l'algèbre, on peut généraliser les résultats à plus de deux événements :

- $$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad -P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad +P(A \cap B \cap C) \end{aligned}$$

- $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ si $A_i \cap A_j = \emptyset \quad \forall i \neq j$

Exemple

On s'intéresse au poids W d'individus masculins

- ☐ A ≡"un individu pèse au moins 55 kg"
- ☐ B ≡"un individu pèse au plus 105 kg"
- ☐ C ≡"un individu pèse au moins 55 kg et au plus 105 kg"

On donne $P(A) = 0.95$ et $P(B) = 0.97$
Que vaut $P(C)$?

Indice : Définissez les événements supplémentaires suivants :

- ☐ A^* ≡"un individu pèse moins de 55 kg"
- ☐ B^* ≡"un individu pèse plus de 105 kg"

Probabilité conditionnelle

Pour évaluer la fréquence d'une maladie bovine et l'efficacité d'un test sanguin, on fait des comptages dans un troupeau de n vaches.

$M \equiv$"la vache est malade" $T \equiv$"le test est positif"
$M^* \equiv$"la vache est saine" $T^* \equiv$"le test est négatif"

	M	M^*	Total
T	$n(M \cap T)$	$n(M^* \cap T)$	$n(T)$
T^*	$n(M \cap T^*)$	$n(M^* \cap T^*)$	$n(T^*)$
Total	$n(M)$	$n(M^*)$	n

Quelle est :

- la fréquence de M quand T est réalisé ?
- la fréquence de T^* quand M est réalisé ?

Probabilité conditionnelle

- Définition de la probabilité conditionnelle $P(A|B)$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad \text{avec } P(B) \neq 0$$

- Elle vérifie les 3 axiomes de base :
  - $P(A|B) \geq 0$
  - $P(\Omega|B) = 1$
  - $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$ si $A_1 \cap A_2 = \emptyset$

Probabilité conditionnelle

- Les probabilités conditionnelles ont des propriétés semblables aux probabilités non conditionnelles :
  - $P(|B) = 0$
  - $P(A|B) \leq 1$
  - $P(A^*|B) = 1 - P(A|B)$
  - $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$

Exercice

Dans un troupeau de 100 vaches, 7 sont malades. On sait qu'un test est positif dans 98 % des cas quand la vache est malade. On prend une vache au hasard. Quelle est la probabilité qu'elle soit malade et que le test soit positif ?

Théorème des probabilités composées

On sait que : $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ $\qquad P(B|A) = \dfrac{P(A \cap B)}{P(A)}$

On a donc : $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

On peut généraliser pour 3 événements :

$$
\begin{aligned}
P(A \cap B \cap C) &= P(A|(B \cap C))P(B \cap C) \\
&= P(A|(B \cap C))P(B|C)P(C)
\end{aligned}
$$

L'intersection d'événements peut s'exprimer comme un produit de probabilités conditionnelles

Théorème des probabilités composées

De manière générale, pour n événements :

$$
\begin{aligned}
&P(A_1 \cap \cdots \cap A_n) \\
\\
&= P(A_1 \cap \cdots \cap A_{n-1})P(A_n|(A_1 \cap \cdots \cap A_{n-1})) \\
&\vdots \\
&= P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \cdots P(A_n|(A_1 \cap \cdots \cap A_{n-1}))
\end{aligned}
$$

Exercice

Dans le troupeau de 100 vaches dont 7 sont malades, on sélectionne n vaches au hasard, qu'on sépare du troupeau. On veut calculer la probabilité qu'aucune des vaches sélectionnées ne soit malade quand $n = 1, 2, 3, \ldots$

Théorème des probabilités totales

Soient des événements $A_1, A_2, \ldots, A_n$ formant une partition.
Soit un événement B quelconque.

On montre alors que :

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)$$

Preuve :

Exercice

Un lot de 21 pièces est issu de trois techniques de fabrication différentes.

Technique n^o	1	2	3
Probabilité de défectuosité	0.01	0.02	0.005
Nombre de pièces produites	7	9	5

Quelle est la probabilité qu'une pièce prise au hasard soit défectueuse ?

Théorème de Bayes

Soient des événements $A_1, A_2, \ldots, A_n$ formant une partition.
Soit un événement B quelconque.

On montre alors que :

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)} \qquad \forall i = 1, \ldots, n$$

Preuve :

Exercice

Un lot de 21 pièces est issu de trois techniques de fabrication différentes.

Technique n^o	1	2	3
Probabilité de défectuosité	0.01	0.02	0.005
Nombre de pièces produites	7	9	5

On prend une pièce au hasard. Sachant qu'elle est défectueuse, quelle est la probabilité qu'elle soit issue de la technique n^o i ?

Théorème de Bayes

- Le théorème de Bayes est aussi appelé le *théorème des causes* :

  On connaît la probabilité $P(B|A_i)$ d'un événement B pour chacune des causes $A_1, \ldots, A_n$ possibles. Si l'on observe que B s'est réalisé, quelle est la probabilité $P(A_i|B)$ que ce soit dû à la cause A_i (avec $i = 1, \ldots, n$) ?

- *Exemples d'applications*
  - systèmes experts
  - filtres bayésien anti-spam (e-mail)

Exercice

On place dans un étang des poissons rouge et gris. Un tiers de ces poissons sont d'origine chinoise. Parmi les poissons chinois, 30 % sont rouges. Au total, 20 % des poissons sont rouges. Un promeneur voit un poisson rouge dans l'étang. Quelle est la probabilité qu'il soit d'origine chinoise ?

Exercice ➡

Dans une université, 10 % des étudiants sont d'origine chinoise. Parmi ceux-là, 90 % ont moins de 30 ans. Il y a 5 % d'étudiants chinois ayant moins de 30 ans et effectuant un doctorat. Vous rencontrez par hasard un étudiant qui dit être chinois et avoir moins de 30 ans. Quelle est votre chance de gagner en pariant qu'il effectue un doctorat ?

Exercice

Deux traitements peuvent être appliqués à une plante malade, avec des probabilités de guérison égales à 90 % et 80 %. Un tiers des plantes reçoivent le premier traitement, les autres reçoivent le second traitement. Si une plante malade est guérie, quelle est la probabilité que le premier traitement ait été appliqué ?

Indépendance

- L'indépendance entre A et B est définie par la propriété :

$$A \perp B \quad \Longleftrightarrow \quad P(A \cap B) = P(A)P(B)$$

L'événement A est indépendant de l'événement B si la (non) réalisation de B n'apporte aucune information sur la probabilité de réalisation de A, et vice-versa :

$$P(A|B) = P(A|B^*) = P(A) \qquad P(B|A) = P(B|A^*) = P(B)$$

- Il s'agit d'une notion fondamentale très importante.

Exercice

On teste la présence d'une maladie dans un troupeau de 100 vaches dont 7 sont malades. Deux possibilités existent quand une vache a été testée:

- ☐ elle est retirée du troupeau et ne sera plus testée
- ☐ elle est replacée dans le troupeau sans marque distinctive

Dans les deux cas, quelle est la probabilité qu'aucune des 7 vaches malades ne soit testée si l'on choisit au hasard 10 vaches parmi les 100 vaches du troupeau ?

Remarques

- ne pas confondre indépendance et incompatibilité !
  - ☐ indépendance : $A \perp B \implies P(A \cap B) = P(A)P(B)$
  - ☐ incompatibilité : $A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$

- l'indépendance peut être généralisée à plus de 2 événements:

$$A \perp B \perp C \quad \iff \begin{cases} P(A \cap B \cap C) = P(A)P(B)P(C) \\ P(A \cap C) = P(A)P(C) \\ P(B \cap C) = P(B)P(C) \\ P(A \cap B) = P(A)P(B) \end{cases}$$

- poser l'hypothèse d'indépendance a d'importantes conséquences.