Probabilités : Introduction et modélisation des situations finies

1. Introduction

L'étude des probabilités a pour objet les situations aléatoires. Cela peut paraître un non-sens, car, pour reprendre la définition de Spinoza du hasard : « une chose est dite contingente, eu égard à un manque de connaissance en nous »¹. Cette vision déterministe est bien l'acception quotidienne que nous nous faisons du hasard, l'on dit d'un événement qu'il tient du hasard lorsque que nous ne savons pas ce qu'il le détermine. Quand je lance une pièce, je dis que le résultat est aléatoire car j'ignore tout ce qui le détermine précisément : la force du lancer, le poids de la pièce, les frottements d'airs sur la pièce etc. Aussi, je suis incapable de prévoir ce que va donner le lancer car l'étendue de ma connaissance est limitée : j'ai affaire un événement aléatoire.

Donc, si l'on s'en tient à cette définition, à quoi bon étudier le hasard, car c'est précisément ce qui n'est pas connaissable ?

Revenons à notre pièce, je suis incapable de deviner à l'avance le résultat d'un lancer, en revanche, je sais que si je lance la pièce, disons 100 fois, je vais avoir une cinquantaine de pile et le reste de face. Je ne suis pas donc pas si ignorant, je ne sais pas ce qu'un événement en particulier donne, mais je sais ce qu'une série d'événements similaires donnent en moyenne. Voilà donc ce que nous allons étudier : ce sont les événements dont un résultat individuel n'est pas prévisible, mais dont je sais ce que je peux en espérer, en moyenne, quand il y en a un grand nombre.

Cette étude est au cœur de nombreux domaines :

En physique quantique (qui étudie les phénomènes microscopique), il est impossible de savoir comment va se comporter une particule, mais l'on peut exhiber des lois sur le comportement d'un grand nombre de ces particules.

En finance de marché, je ne peux savoir si un individu va effectuer une transaction un jour donnée mais je peux déterminer combien de transaction vont être effectuées en moyenne par un grand nombre d'utilisateurs un jour donné.

La Française des Jeux ne sait pas quel joueur va gagner, mais elle sait que si un certain nombre de joueur jouent, elle peut espérer un gagnant. Préciser tout cela avec des calculs de probabilités permet de fixer une cagnotte et un prix du ticket, pour que la FDJ soit rentable.

2. Le formalisme probabiliste mathématique

Au cœur de notre étude se situe donc la notion **d'expérience aléatoire**. C'est une expérience, qu'on suppose répliquable, et dont les résultats, qu'on appelle les **issues**, ne sont pas toujours les mêmes à chaque itération de l'expérience.

On note l'ensemble des issues Ω : **l'univers**.

Dans ce cours, il y a un **nombre fini** d'issues.

Les **événements** sont les divers assemblages d'issues, formellement ce sont les parties de Ω , donc l'ensemble des événements est noté $\mathcal{P}(\Omega)$.

On dit qu'un **événement est réalisé** si une de ses issues se réalise.

Exemple 1 : Je lance un dé à 6 faces. Il y a donc 6 issues : $\Omega = \{1,2,3,4,5,6\}$. Un événement est, par exemple, $A = \ll \text{avoir un numéro pair} \gg$. Donc $A = \{2,4,6\} \in \mathcal{P}(\Omega)$. Si le dé donne 2 alors A est réalisé.

Les événements étant des ensembles, on peut leur appliquer les opérations ensemblistes vues en début d'année : union, intersection, complémentaire, privé de etc.

¹ Éthique, Livre I

Définition : Soit A un événement d'un univers Ω . Le **complémentaire (ou contraire)** de A , noté $\bar{A}$ contient toutes les issues n'étant pas dans A .

Dans l'exemple 1, $\bar{A} = \text{« obtenir un numéro impair »} = \{1,3,5\}$.

Définition : On dit que deux événements A et B sont **incompatibles** s'ils n'ont pas d'issues en commun : $A \cap B = \emptyset$.

Concrètement, cela signifie que A et B ne peuvent se réaliser en même temps.

Dans l'exemple 1 : A et l'événement $C = \text{« obtenir un multiple de 5 »} = \{5\}$ sont incompatibles.

Un événement et son contraire sont toujours incompatible.

Définition : Soient $(A_1, \dots, A_n)$ une famille d'événements ($n \geq 1$). On dit que cette famille est un **système complet d'événement** si :

- Les événements sont 2 à 2 incompatibles
- $\bigcup_{i=1}^n A_i = \Omega$

En d'autres termes, un système complet d'événement est une partition de l'univers. On fait une espèce de disjonction de cas.

Remarque : Pour tout événement A , la famille $(A, \bar{A})$ est toujours un système complet par définition du complémentaire. Cela signifie que soit A est réalisé soit $\bar{A}$ est réalisé.

Exemple : Si je sélectionne un élève au hasard dans la classe, et que je note les événements $E = \text{« l'élève fait espagnol »}$; $A = \text{« l'élève fait arabe »}$; $D = \text{« l'élève fait allemand »}$ et $C = \text{« l'élève fait chinois »}$. Alors (E, A, D, C) forme un système complet d'événement, en effet un élève ne peut faire deux de ces langues en même temps (c'est l'incompatibilité des événements) et un élève fait au moins une de ces langues (c'est le fait que la réunion vaut l'univers).

On n'a toujours pas parlé de probabilité ! Une probabilité est une façon de mesurer les événements, en leur assignant à chacun un nombre entre 0 et 1.

Définition : Une probabilité sur un univers Ω est une application $P : \mathcal{P}(\Omega) \rightarrow [0,1]$ vérifiant :

- $P(\Omega) = 1$
- Pour tous événements A et B incompatibles : $P(A \cup B) = P(A) + P(B)$

Remarque : Le point 2 est équivalent à (par récurrence) à :

- Pour toute famille d'événements 2 à 2 incompatibles $(A_1, \dots, A_n)$:
 $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.
Cette propriété s'appelle la σ - *additivité*.

Propriétés : Soit P une probabilité, A , B et C des événements. On a :

- $P(\bar{A}) = 1 - P(A)$
- $P(\emptyset) = 0$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $A \subset B \Rightarrow P(A) \leq P(B)$
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Propriété (Formule des Probabilités Totales) :

Étant donné un système complet d'événement $(A_1, \dots, A_n)$ et B un événement. On peut décomposer la probabilité de B comme ceci :

$$P(B) = \sum_{i=1}^n P(B \cap A_i)$$

Remarque 1 : En considérant un événement A et le système complet $(A, \bar{A})$ l'on obtient :

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

Ce qui traduit le fait que « Soit B et A se réalisent soit B et non A se réalisent »

Nul besoin de symbole somme si le système complet ne contient que 2 ou 3 événements !

Remarque 2 : C'est la formule la plus importante du cours de probabilités discrètes. Elle traduit une disjonction de cas selon les différents événements $(A_1, \dots, A_n)$. On l'utilise lorsqu'il est difficile de calculer la probabilité de B directement et que l'on sait la calculer si l'on fixe un événement A_i .

Définition : On est en situation d'équiprobabilité (pour une probabilité P) si chaque issue de l'univers Ω a autant de chance de se réaliser que les autres.

Ainsi, comme la somme des probabilités de chaque issue vaut 1, pour toute issue ω_i :

$$P(\{\omega_i\}) = \frac{1}{\text{Card}(\Omega)} = \frac{1}{n}$$

Exemple : Un lancer de pièce équilibrée, un lancer de dé équilibré, le résultat de 3 lancers de dés équilibrés. Bref, quand les issues sont équilibrées.

Remarque : En pratique, on choisit de modéliser une situation par de l'équiprobabilité, lorsqu'il n'y a pas de raison qu'une issue ait une probabilité différente des autres (pensez à un lancer de pièce, on va dire qu'il y a une chance sur deux car on ne voit pas pourquoi un coté aurait plus de chance d'apparaître que l'autre).

Attention : tout univers ne se modélise pas avec de l'équiprobabilité, cela dépend de ce qu'on considère être les issues.

Par exemple, prenons le résultat du Loto et considérons l'univers $\Omega = \{\text{je gagne, je perds}\}$. Évidemment il n'y a pas d'équiprobabilité.

Il y en a si je considère les issues comme étant les résultats du Loto (cela fait donc des millions d'issues et donc une chance sur des millions de gagner si on joue une seule fois).

Proposition : Étant donné un événement A, en cas d'équiprobabilité :

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{\text{Nombre de cas favorables pour A}}{\text{Nombre de cas possibles}}$$

Exemple : Je lance deux dés équilibrés à 6 faces. Il y a 36 issues (les couples de résultats) et il y a d'équiprobabilité. Si je note A = « obtenir 2 fois le même numéro », donc $A = \{(1,1); (2,2); (3,3); (4,4); (5,5); (6,6)\}$, ainsi :

$$P(A) = \frac{\text{Nombre de cas favorables pour A}}{\text{Nombre de cas possibles}} = \frac{6}{36} = \frac{1}{6}$$

On peut aussi se dire qu'avoir deux fois le même numéro c'est avoir au second lancer le numéro du premier, donc 1 chance sur 6.

Il n'y a jamais qu'une seule façon d'obtenir un résultat en probabilité. Multiplier les points de vue sur un problème est très riche et permet de fournir des solutions parfois plus rapides.

3. Probabilités, fréquence et simulation informatique

Nous venons de voir que la probabilité est une modélisation. Comme toute modélisation, sa valeur se mesure à l'aune de son efficacité pratique, de sa capacité prédictive.

Cette efficacité s'évalue par le calcul des fréquences des événements. En effet, pour que la probabilité d'un événement soit légitime, il faut absolument que cette probabilité coïncide avec la fréquence d'apparition de cet événement.

Prenons un exemple, si je lance 100 fois une pièce équilibrée et que je calcule la fréquence d'apparition de l'événement $A = \text{« la pièce donne pile »}$. Je vais obtenir une fréquence proche de $50\%=0,5$, c'est-à-dire la probabilité théorique de l'événement A .

Ces considérations se résument en la Loi des Grands Nombres qui sera vue plus tard et qui affirme que, pour un grand nombre de répétitions, la fréquence observée de l'événement A est aussi proche que l'on veut de sa probabilité. Voyons ce que cela donne avec l'ordinateur.

L'outil informatique permet de simuler du hasard, c'est-à-dire, de produire des expériences aléatoires, c'est-à-dire en pratique générer des nombres de manière aléatoires². Par exemple, afficher un nombre entre 1 et 10 de manière équiprobable.

Cela peut paraître fort étonnant car le principe même d'un algorithme est d'exécuter systématiquement les mêmes opérations. En générant un nombre au hasard, ce qui se fait en Python à l'aide de la commande `rd.randint(1,11)` pour des entiers entre 1 et 10, le programme va donc générer à chaque nouvelle exécution un résultat différent. Il suffit de s'imaginer qu'il y a un lutin dans la machine qui fait des lancers de pièce et qui affiche ses résultats.

Considérons le programme suivant qui génère aléatoirement 100 fois un entier entre 0 et 1 et qui calcule et affiche la fréquence d'apparitions de l'entier 1.

```
N=100
S=0
for k in range(0,N):
    x=rd.randint(0,2)
    S=S+x
print(S/N)
```

Il s'affiche à l'écran : 0.54.

C'est une illustration de la Loi des Grands Nombres, la fréquence pour un grand nombre (ici 100) de répétitions de l'expérience est proche de la probabilité de l'événement « obtenir 1 » qui est de 0,5.

En remplaçant N par 1000, il s'affiche 0.508. On se rapproche de 0,5 !

² Ce sont les Random Generated Numbers *RNG* que connaissent bien les joueurs de jeux vidéo.