Chapitre 1

Probabilité

1.1 Introduction

Lorsqu'on lance une pièce de monnaie, on ne sait pas avec certitude lequel des deux résultats possibles surviendra : pile (P) ou face (F). Cependant, si on répète l'expérience plusieurs fois et que la pièce est bien équilibrée, on s'attend à ce que la proportion de fois qu'on obtienne F soit près de $1/2$. En fait, on s'attend à observer une proportion de F d'autant plus près de $1/2$ avec d'autant plus de certitude que le nombre de lancers est grand. On s'attend même à ce que cette proportion tende vers $1/2$ lorsque le nombre de lancers tend vers l'infini. C'est la magie des grands nombres qui transforme une incertitude en certitude. Vous pouvez vérifier ce phénomène en lançant une pièce de monnaie 100 fois et en notant la proportion de F après chaque dizaine de fois.

Remarquons que lancer une pièce de monnaie bien équilibrée une fois et observer le résultat revient à tirer une répétition de cette expérience au hasard parmi une infinité de répétitions virtuelles dont la moitié donnent F comme résultat et l'autre moitié P. La probabilité d'observer F est alors $1/2$. Il s'agit ici d'une probabilité *a priori* qui mesure le degré de certitude d'observer F. La valeur de cette probabilité repose cependant sur une hypothèse, soit l'hypothèse de symétrie entre les deux résultats possibles.

Dans le cas d'un nombre fini quelconque de résultats possibles, l'hypothèse de symétrie mène à la définition classique de la probabilité. Celle-ci a été introduite au 17^{e} siècle principalement en lien avec les jeux de hasard.

Mentionnons notamment Blaise Pascal (1623-1662) et Pierre de Fermat (1601-1665) qui se demandent dans une correspondance datant de 1654 comment partager les mises au jeu entre deux joueurs qui avaient convenu de déclarer vainqueur le premier qui gagnerait un nombre donné de parties de pur hasard et qui doivent interrompre la série de parties avant que le vainqueur ne soit déclaré. Un partage équitable repose sur la détermination d'événements équiprobables. À peu près au même moment, soit en 1657, Christiaan Huygens (1629-1695) publie *De ratiociniis in ludo aleae* (Raisonnements sur les jeux de dés) dans lequel il traite du prix honnête à payer pour jouer à un jeu de hasard, ce qui introduit la notion d'espérance mathématique calculée à partir des probabilités des issues possibles du jeu et des gains qui leur sont associées.

De façon plus générale, on peut s'intéresser à la probabilité de réalisation d'un événement suite à une expérience aléatoire, c'est-à-dire une expérience dont on ne peut pas prévoir le résultat avec certitude du fait qu'elle est soumise au hasard, et représenter cette probabilité par un nombre p compris entre 0 et 1. Même si ce nombre est inconnu, il peut être approché par la fréquence de réalisation de cet événement en répétant l'expérience aléatoire un grand nombre de fois. Les répétitions doivent cependant se dérouler indépendamment dans les mêmes conditions. Tout cela est précisé dans la loi des grands nombres dont la première version est donnée par Jakob Bernoulli (1654-1705) dans *Ars conjectandi* publié à titre posthume en 1713. C'est cette loi qui justifie l'interprétation fréquentiste de la probabilité.

Il faut attendre plusieurs années avant que ne soit décrite la distribution en forme de cloche, dite normale, des fréquences de réalisation d'un événement, et plus généralement des moyennes de valeurs observées d'une variable aléatoire, sous l'hypothèse d'un grand nombre d'observations faites indépendamment dans les mêmes conditions. On doit cette description dans sa forme la plus simple à Abraham de Moivre (1667-1754) en 1733, et dans des formes de plus en plus élaborées à Pierre-Simon de Laplace (1749-1827) qui culminent avec la publication en 1812 de l'ouvrage *Théorie analytique des probabilités*. Le résultat a été par la suite généralisé en affaiblissant les conditions au maximum, et plus tard étendu à d'autres situations.

Bien qu'il soit toujours possible en théorie de considérer la probabilité de réalisation d'un événement, il n'est pas toujours possible de reproduire les mêmes conditions d'observation. On peut penser, par exemple, à une affirmation sur le temps qu'il fera demain ou sur le changement journalier de la valeur d'un actif financier. Il y a néanmoins un intérêt à considérer la probabilité que l'affirmation soit vraie et à interpréter cette probabilité comme si on pouvait effectuer un grand nombre d'observations indépendamment dans les mêmes conditions. Encore faut-il préciser pour quelle collection d'événements une probabilité est bien définie et quelles sont les règles de calcul qui s'appliquent. De telles considérations ont mené Andreï Nikolaïevitch Kolmogorov (1903-1987) à une axiomatisation des probabilités présentée dans l'ouvrage *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Fondements de la théorie des probabilités) publié en 1933. C'est dans ce cadre que se développe encore aujourd'hui la théorie des probabilités.

La théorie des probabilités permet de décrire un grand nombre de phénomènes. Elle est à la base de plusieurs domaines connexes, notamment la statistique, l'actuariat et la finance mathématique, et on trouve des applications dans pratiquement toutes les sciences, que ce soient les sciences de la nature, les sciences sociales ou la science des données.

1.2 Définition classique d'une probabilité

Considérons une *expérience aléatoire*, donc une expérience dont l'issue n'est pas certaine, qui donne lieu à un nombre fini N de résultats possibles. On peut penser, par exemple, au lancer d'une pièce de monnaie ou d'un dé.

Un *événement* A est une affirmation sur le résultat d'une expérience aléatoire. Les résultats pour lesquels un événement A se réalise, appelés *résultats favorables*, forment un sous-ensemble de tous les *résultats possibles*. Ce sous-ensemble est également représenté par A.

Supposons maintenant qu'un événement A se réalise pour n résultats parmi N qui ont des chances égales de se produire. La *probabilité* de A est alors donnée par n/N. Autrement dit, on définit

$$\mathbb{P}(A) = \frac{\text{nombre de résultats favorables}}{\text{nombre de résultats possibles}}.$$

Cette définition revient à dire que les résultats possibles sont équiprobables, en l'occurrence de même probabilité $1/N$. C'est un *modèle uniforme discret*. De plus, la probabilité est additive, dans le sens que la probabilité d'un événement A est égale à la somme des probabilités des résultats pour lesquels l'événement A se réalise. C'est la définition classique de la probabilité.

L'équiprobabilité de résultats possibles repose sur une hypothèse de symétrie. Cette hypothèse découle d'un *principe d'indifférence* tel qu'énoncé par Pierre-Simon de Laplace en 1814 dans *Essai philosophique sur les probabilités* : en cas d'incertitude de même degré sur la survenance d'un certain nombre de résultats possibles, les probabilités de ces résultats sont égales. C'est l'hypothèse qui est sous-entendue par défaut lorsqu'on dit qu'un résultat survient au hasard.

Dans l'exemple du dé, l'hypothèse de symétrie revient à supposer que le dé est non pipé. De plus, obtenir un nombre pair de points est un événement qui se réalise si et seulement si on obtient 2, 4 ou 6 points, et donc, par additivité, la probabilité de cet événement est 3/6, soit 3 fois 1/6.

Avec la définition classique, calculer la probabilité d'un événement revient d'abord à déterminer quels sont les résultats possibles équiprobables, puis à les dénombrer, et enfin à dénombrer les résultats favorables à l'événement considéré parmi tous les résultats possibles. Cette procédure peut paraître simple, mais des difficultés surviennent dans la détermination des résultats possibles qui sont équiprobables et le dénombrement de résultats nombreux.

1.2.1 Problème des partis

Deux joueurs misent chacun la même somme d'argent pour avoir le droit de jouer à un jeu de pur hasard qui se déroule en trois manches gagnantes (ou plus) mais le jeu est interrompu avant que l'un des deux ait gagné trois manches. Les deux joueurs sont alors d'accord pour partager la mise. Le problème consiste alors à savoir quel est le partage juste ou équitable.

C'est ainsi que Pierre Fermat pose le *problème des partis* à Blaise Pascal dans sa lettre datée du 29 juillet 1654. Remarquons d'abord que, pour être équitable, le partage doit se faire selon les probabilités qu'ont chacun des deux joueurs d'être déclaré vainqueur au moment de l'interruption du jeu. Le problème revient donc à déterminer ces probabilités. La difficulté ici est de déterminer des résultats élémentaires équiprobables à partir desquels ces probabilités peuvent être calculées.

Supposons, par exemple, que les joueurs sont A et B et que le jeu est interrompu alors que A ait gagné 2 manches et B aucune. Pour solutionner le problème du partage équitable de la mise dans cette situation, on peut imaginer que le jeu se poursuit après l'interruption, et ce jusqu'à la 5$^{\text{e}}$ manche. Les gagnants des 3$^{\text{e}}$, 4$^{\text{e}}$ et 5$^{\text{e}}$ manches peuvent être dans l'ordre AAA, AAB, ABA, BAA, ABB, BAB, BBA ou BBB, et ces 8 résultats possibles sont équiprobables par symétrie. Dans les 7 premiers cas, c'est A qui est déclaré le vainqueur, et dans le dernier cas, c'est B. On a donc

$$\mathbb{P}(A \text{ vainqueur}) = \frac{7}{8}, \quad \mathbb{P}(B \text{ vainqueur}) = \frac{1}{8}.$$

Vous pouvez vous amuser à calculer de même la probabilité qu'un joueur qui mène 2 parties à une dans un 4 de 7 contre un adversaire de même force ne soit pas le premier des deux à gagner 4 parties. Le problème général avec un nombre quelconque de parties à gagner pour être déclaré vainqueur et une interruption après des nombres quelconques de parties gagnées par A et par B se résout de façon analogue. Il pose cependant des difficultés de calcul lorsque les nombres sont grands.

1.2.2 Problème du scrutin

À l'occasion d'un scrutin opposant deux candidats A et B, le candidat A reçoit a voix et le candidat B, b voix, a étant supérieur à b. On demande quelle est la probabilité pour que le candidat A soit en avance dans le décompte des voix tout le long de leur dépouillement.

Ce *problème du scrutin* fut proposé par Joseph Bertrand (1822-1900) en 1887. Il peut se résoudre à partir de quelques remarques. Tout d'abord, pour que A soit en avance tout au long du dépouillement dans un ordre donné, ce dépouillement doit nécessairement commencer par une première voix pour A. Par la suite, la première voix pour B doit pouvoir être annulée par une voix précédente pour A autre que la première. En retirant ces deux voix, cela doit aussi être possible pour la seconde voix pour B, et de même pour toutes les voix suivantes pour le candidat B.

Pour calculer la probabilité de l'événement considéré, on dispose d'abord les $a + b$ voix en cercle. Puis, on effectue le dépouillement dans le sens des aiguilles d'une montre à partir de l'une des $a+b$ voix. Tous ces dépouillements possibles sont supposés équiprobables. Pour déterminer ceux avec A en avance tout au long du dépouillement, on dispose toujours les $a + b$ voix en cercle, puis on retire successivement toutes les voix pour A et pour B qui se suivent dans le sens des aiguilles d'une montre. On retire ainsi b paires de voix AB. Il reste donc $a + b - 2b = a - b$ voix pour A. Les dépouillements dans le sens des aiguilles d'une montre à partir de ces $a - b$ voix sont ceux avec un nombre accumulé de voix pour A toujours supérieur au nombre accumulé de voix pour B. Le nombre de dépouillements favorables à cet événement est le même pour toute disposition initiale des voix. La probabilité cherchée est donc

$$\mathbb{P}(A \text{ en avance tout au long du dépouillement}) = \frac{a-b}{a+b}.$$

Vous pouvez vérifier cette formule dans le cas $a = 10$ et $b = 6$ en disposant en cercle dans un ordre quelconque les 10 voix pour A et les 6 voix pour B, puis en retirant successivement toutes les voix consécutives pour A et B dans le sens des aiguilles d'une montre. Vous constaterez que les dépouillements dans le sens des aiguilles d'une montre qui commencent par les voix restantes pour A sont ceux avec A en avance tout au long du dépouillement.

1.3 Règles de dénombrement

Le principal défi posé par la définition classique de la probabilité est de dénombrer des résultats. Ce défi peut être relever avec succès à partir de deux principes fondamentaux et de quelques formules de base en analyse combinatoire.

1.3.1 Principe de multiplication

Le nombre total de résultats possibles de r expériences ordonnées avec n_1 résultats possibles pour la première, puis n_2 pour la deuxième, et ainsi de suite jusqu'à n_r pour la r-ième est

$$n = n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i.$$

C'est le *principe de multiplication.* Il est à remarquer que les résultats possibles pour chaque expérience peuvent dépendre des résultats des expériences précédentes. Ce qui est important, c'est que leur nombre ne dépende pas des résultats antérieurs.

Beaucoup de problèmes de dénombrement peuvent être résolus en appliquant le principe de multiplication. Il suffit de concevoir une suite ordonnée d'expériences avec des résultats possibles en nombre fixe à chaque étape pour décrire tous les résultats dans leur ensemble.

Exemple. On considère les anniversaires de naissance de n personnes nées une année non bissextile. Il y a *a priori* 365 possibilités pour chacune des n personnes dans un ordre quelconque, d'où 365 multiplié n fois pour le nombre total de possibilités. D'autre part, pour qu'aucune personne n'ait son anniversaire le même jour qu'une autre, il y a toujours 365 jours possibles pour l'anniversaire de la première, mais 364 jours possibles pour l'anniversaire de la seconde (en retirant le jour anniversaire de la première), et ainsi de suite jusqu'à $365 - (n-1) = 365 - n + 1$ jours possibles pour l'anniversaire de la n-ième personne. En supposant toutes les possibilités *a priori* équiprobables, on a alors

$$\mathbb{P}(n \text{ anniversaires tous différents}) = \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365 \times 365 \times \cdots \times 365}.$$

En effectuant le calcul de cette probabilité pour différentes valeurs de n, on constate qu'elle est inférieure à $1/2$ dès que $n \geqslant 23$. Vous pouvez vérifier cela à l'aide d'une calculatrice.

1.3.2 Nombre de permutations

Une *permutation* de n éléments est une disposition ordonnée des n éléments, et il y en a un nombre total donné par la *factorielle* de n, soit

$$n! = n \times (n-1) \times \cdots \times 1.$$

En admettant le principe de multiplication, la formule pour le nombre de permutations de n éléments est immédiate. En effet, il suffit de choisir un premier élément pour lequel il y a n possibilités, puis un second pour lequel il reste $n-1$ possibilités, et ainsi de suite jusqu'au n-ième élément pour lequel il ne reste qu'une seule possibilité.

Exemple. Supposons qu'une personne veuille appeler à tour de rôle ses trois meilleurs amis, disons A, B et C, pour leur annoncer une grande nouvelle. Elle peut d'abord commencer par n'importe lequel des trois, puis poursuivre avec n'importe lequel des deux autres, et finalement terminer avec le seul qui reste. Il y a en tout $3 \times 2 \times 1 = 6$ ordres d'appel possibles.

Remarque. Le nombre $n!$ devient difficile à calculer lorsque n devient grand. Cela était particulièrement vrai avant l'avènement des ordinateurs. Il existe heureusement une formule qui permet d'approcher sa valeur, soit

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n,$$

qui est connue sous le nom de *formule de Stirling.* La constante π, parfois appelée *constante d'Archimède*, est approchée par 3,141 592 653. Quant à la constante e, appelée *nombre d'Euler* ou *constante de Néper*, sa valeur est approchée par 2,718 281 828. Le symbole $\sim$ signifie que le rapport des termes à gauche et à droite tend vers 1 lorsque n croît. Ainsi, pour $n = 5$ et 10 la valeur exacte de $n!$ est donnée par 120 et 3 628 800, respectivement, et la valeur approchée par 118,02 et 3 598 696, respectivement, pour un rapport égal à 1,016 777 et 1,008 365, respectivement. Vous pouvez effectuer vous-même les calculs pour $n = 20$ à l'aide d'une calculatrice et comparer les valeurs exacte et approchée. On voit cependant que la valeur approchée de $n!$ est très près de la valeur exacte même pour de petites valeurs de n. La formule de Stirling permet non seulement d'approcher la factorielle de tout nombre entier n, mais aussi de montrer des résultats limites pour des fonctions de $n!$.

1.3.3 Nombre de combinaisons

Une *combinaison* de k éléments parmi n est un sous-ensemble de k éléments parmi les n, et il y en a un nombre total donné par

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

avec la convention $0! = 1$.

On peut obtenir le nombre de combinaisons de k éléments parmi n à partir du nombre de dispositions ordonnées de k éléments parmi n, appelées *arrangements*. Par le principe de multiplication, ce nombre est donné par

$$n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!},$$

et ce en choisissant un premier élément pour lequel il y a n possibilités, puis un second pour lequel il reste $n-1$ possibilités, et ainsi de suite jusqu'au k-ième élément pour lequel il ne reste que $n-(k-1) = n-k+1$ possibilités. Or, il est aussi donné par le nombre de sous-ensembles de k éléments parmi n multiplié par le nombre de dispositions ordonnées de ces k éléments, d'où

$$\binom{n}{k} \times k! = \frac{n!}{(n-k)!}.$$

On obtient la formule pour le nombre de combinaisons de k éléments parmi n en divisant par $k!$.

À remarquer qu'on a

$$\binom{n}{k} = \binom{n}{n-k}.$$

En effet, le nombre de sous-ensembles de k éléments d'un ensemble de n éléments correspond au nombre de sous-ensembles dont le complémentaire possède $n-k$ éléments.

Exemple. On considère le nombre de comités possibles composés d'un président, d'un trésorier, d'un secrétaire et d'un membre, sans cumul de fonctions, tous choisis dans un groupe de 20 personnes. Ce nombre est donné par

$$20 \times 19 \times 18 \times 17 = \binom{20}{4} \times 4!.$$

En effet, on peut choisir d'abord un président, puis un trésorier, un secrétaire et enfin un membre, ou encore choisir d'abord quatre personnes, puis leur assigner des fonctions. D'autre part, la probabilité qu'il y ait parité des genres au sein du comité, si cette parité existe dans le groupe et si tous les comités sont *a priori* équiprobables, est

$$\frac{\binom{10}{2} \times \binom{10}{2}}{\binom{20}{4}} = \frac{10!10!(20-4)!4!}{(10-2)!2!(10-2)!2!20!}.$$