

C'est quoi, une probabilité ?

C'est quoi, une probabilité ?

C'est une fonction P définie sur un univers Ω et à valeurs dans $[0; 1]$, telle que la somme de ses valeurs pour tous les éléments de l'univers soit égale à 1.

Euh...

Ah oui, pardon, il faut que je précise que dans cette définition, l'univers est fini. Il s'agit de probabilités discrètes.

Euh...

Tu comprendras quand tu seras plus grande ...

Non, mais en vrai, c'est quoi une probabilité ?

I Premier point de vue : une fréquence

Le concept de probabilité cherche à mesurer les chances qu'on a de gagner à un jeu de hasard. Si je lance une pièce de monnaie, quelles chances as-tu d'obtenir pile ?

En trichant ou pas ? Je peux utiliser ma pièce truquée ?

Non, sans tricher, bien sûr.

Eh bien, j'ai une chance sur deux.

Oui, mais qu'est-ce que ça veut dire, et pourquoi est-ce comme ça ?

Ca veut dire qu'il y a deux résultats possibles (pile ou face) et que je m'intéresse à un seul d'entre eux (pile). Et c'est comme ça parce que les chances d'avoir chacun sont égales puisque la pièce n'est pas truquée. Donc une chance (pile) sur deux (pile ou face).

Oui, c'est effectivement l'idée de départ. Mais il faut arriver à justifier d'une manière plus scientifique cette égalité des chances. Quelle expérience peut-on faire qui mettrait en évidence cette égalité des chances ?

On lance la pièce un grand nombre de fois (mettons 100 fois) et on compte le nombre de fois où on obtient pile. Normalement, on devrait trouver 50. Enfin, peut-être pas tout le temps, mais ça ne devrait pas être très loin de 50.

C'est un peu l'idée, mais ce n'est pas si simple que ça, parce qu'il s'agit d'une expérience aléatoire (dont le résultat dépend du hasard). En lançant une pièce 100 fois, on peut très bien obtenir 80 pile.

C'est sans doute parce que 100 est trop petit. Peut-être faut-il lancer la pièce plus de fois, par exemple 1000 fois

Oui, mais même en la lançant 1000 fois, on peut obtenir 800 fois pile.

J'ai du mal à le croire. Normalement, plus on lance la pièce, plus le nombre de pile et de face tend à s'équilibrer.

D'une certaine manière, oui, mais il va falloir préciser ce que ça veut dire. On peut quand même obtenir 800 fois pile. C'est le propre du hasard : les choses peuvent se produire, les catastrophes aussi bien que les coups de chance, même si leurs chances sont faibles. Le propre du hasard, c'est l'incertitude : on parie, mais on peut gagner ou perdre, que les chances soient grandes ou pas. C'est effectivement très déroutant, mais c'est quasiment la définition du hasard : c'est ce qui est imprévisible, malgré toutes les idées a priori qu'on peut avoir. Si on peut prévoir quelque chose avec certitude, alors ce n'est plus du hasard.

Quoi qu'il en soit, ce qui compte dans ton expérience, ce n'est pas tellement le nombre exact de pile (ça ne sera pas le même avec 100 lancers ou 1000), c'est la proportion des pile, c'est-à-dire le rapport du nombre de pile au nombre de lancers. On appelle ce nombre la *fréquence* de réalisation de l'événement « pile ». Ce terme de *fréquence* est d'ailleurs souvent employé aussi en statistiques, quand on étudie de grandes quantités de données.

Lorsque le nombre de lancers est très grand, quelle fréquence s'attend-on à trouver à peu près ?

Par exemple $\frac{50}{100}$, c'est-à-dire 0,5, c'est-à-dire $\frac{1}{2}$

Oui, et c'est cette fréquence « idéale » qu'on appelle la *probabilité* de l'événement pile. On s'attend normalement à ce que, plus le nombre de lancers est grand, plus la fréquence réelle se rapproche de la probabilité.

La probabilité serait la limite de la fréquence lorsque le nombre de lancers tend vers l'infini ?

Oui, on peut dire ça, mais la traduction mathématique précise de cette idée est d'une redoutable difficulté, parce qu'elle est assez abstraite. On l'appelle souvent la « loi des grands nombres ».

Le problème, c'est qu'aucune expérience réelle ne peut atteindre l'infini. Et donc le choix de la valeur théorique de 0,5 vient en fait de deux choses : d'abord des expériences réelles suggèrent des valeurs voisines de 0,5, et ensuite un principe a priori d'égalité des chances suggère aussi cette valeur de 0,5 : il n'y a pas de raison a priori pour qu'un des côtés de la pièce tombe plus souvent que l'autre. Donc il serait étonnant et peu justifiable de choisir des probabilités différentes.

1) Les propriétés des probabilités

Le fait d'assimiler une probabilité à une fréquence impose certaines propriétés sur les probabilités.

Valeurs possibles d'une probabilité

D'abord une fréquence se calcule en divisant le nombre de fois où on a obtenu pile par le nombre de lancers, par exemple $\frac{42}{100}$. Donc quelles sont les valeurs possibles d'une fréquence, et les valeurs impossibles ?

Une fréquence est forcément inférieure ou égale à 1 (le nombre de pile est forcément inférieur ou égal au nombre de lancers).

Oui, mais ce n'est pas tout. Quel est le signe d'une fréquence ?

Ah oui, c'est forcément positif puisque c'est le rapport de deux nombres entiers.

Oui. Plus précisément « positif ou nul » et « nombres entiers naturels » (avec de plus un dénominateur qui ne peut pas être nul).

Donc une probabilité est forcément une fraction dans $[0; 1]$?

Non, pas tout à fait. C'est bien une valeur dans $[0; 1]$, mais ce n'est pas forcément une fraction.

Mais pourquoi ? On a bien dit qu'une fréquence était une fraction !

Oui, mais on n'a pas dit qu'une probabilité était une fréquence, seulement que c'était « analogue » à une fréquence. En fait, une probabilité est une *limite* d'une suite de fréquences. Le problème, c'est que la limite d'une suite de trucs n'est pas forcément un truc.

Par exemple, la limite d'une suite de fractions n'est pas forcément une fraction. On peut fabriquer une suite de fractions dont la limite est $\sqrt{2}$, qui n'est pas une fraction. Une telle suite peut être définie par $u_0 = 1$ et $u_{n+1} = \frac{1}{2} \left(u_n + \frac{2}{u_n} \right)$. Pour tout n , u_n est bien une fraction (d'après les calculs sur les fractions et les nombres entiers), mais sa limite n'en est pas une.

De même, la limite d'une suite de nombres strictement positifs n'est pas forcément strictement positive. Par exemple avec $u_n = \frac{1}{n}$ pour $n \geq 1$. Sa limite est 0 et pourtant $u_n > 0$ pour tout n .

Somme des probabilités

Quand on lance une pièce 100 fois, si on obtient x fois pile et y fois face, on a nécessairement $x + y = 100$, puisque tout résultat est forcément soit pile soit face.

Les fréquences sont $p = \frac{x}{100}$ et $f = \frac{y}{100}$. Elles vérifient donc $p + f = \frac{x + y}{100} = \frac{100}{100} = 1$. Cette propriété est encore vraie s'il y a plus de deux résultats possibles (par exemple, pour un dé, il y a 6 résultats possibles) :

la somme des probabilités des résultats possibles est égale à 1.

C'est comme dans un sondage : la somme des réponses (y compris les réponses « je ne réponds pas ») doit faire 100% !

Exactement. 100% veut d'ailleurs dire $\frac{100}{100}$, c'est-à-dire 1.

Du point de vue mathématique, lorsqu'une expérience a un ensemble U de résultats possibles, définir des probabilités relativement à cette expérience revient donc à choisir pour chaque résultat une valeur réelle dans $[0; 1]$, de telle sorte que la somme des probabilités des résultats soit égale à 1.

Oui, mais on ne peut pas choisir les probabilités n'importe comment ! Il faut quand même que les probabilités soient les limites des fréquences réelles.

Il y a ici une difficulté : la différence entre, d'une part, la modélisation mathématique d'une expérience réelle et, d'autre part, les concepts mathématiques qui constituent le modèle.

2) La modélisation

Dans une démarche scientifique, on part d'une expérience réelle et on en propose un modèle mathématique qui est censé la représenter fidèlement. Le problème, c'est qu'on ne peut pas démontrer mathématiquement qu'un modèle mathématique représente bien une situation de la réalité. C'est le problème de toutes les sciences de la nature. Or les mathématiques ne sont pas une science de la nature. Elles sont purement conceptuelles. Les outils de démonstration qu'elles offrent permettent de démontrer que certaines propriétés conceptuelles en impliquent d'autres, mais pas qu'elles sont adaptées à la représentation des phénomènes naturels. On peut essayer de « vérifier » ou de « confirmer » cette adaptation, mais cela ne peut pas être par une démonstration mathématique. Cela se fait en général par des expériences ou des raisonnements de type physique.

Ensuite, lorsqu'on a choisi un modèle (en espérant qu'il soit adapté, et en ayant si possible confirmé cette adaptation par des expériences), les raisonnements qu'on fait sur le modèle sont des raisonnements mathématiques.

C'est pour cela que les théories et les concepts mathématiques évacuent toute référence à la « nature » ou à la « réalité ». Elles posent des définitions et des axiomes de nature conceptuelle, et partent de là pour prouver des propriétés, en se limitant à utiliser des raisonnements de type mathématique.

Cela explique pourquoi, pour un mathématicien, la définition d'une probabilité comme on l'a donnée au début suffit :

Une probabilité (ou plutôt une loi de probabilité) est une fonction P définie sur un univers Ω et à valeurs dans $[0; 1]$, telle que la somme de ses valeurs pour tous les éléments de l'univers soit égale à 1.

On comprend bien sûr d'où viennent ces idées : elles viennent des fréquences. Mais une fois cette définition posée, le concept vit sa vie propre dans le monde des mathématiques. C'est pour cela qu'il y a plusieurs lois de probabilité possibles sur un même univers. Il y en a même une infinité. Savoir si ces lois « représentent » fidèlement une situation réelle est un autre problème, en dehors des mathématiques (mais rien n'empêche un

mathématicien d'avoir des idées en dehors des mathématiques, cela serait même plutôt recommandé).

Mais alors, comment puis-je justifier mon choix de modèle sur une copie ? Pendant une contrôle, je ne peux pas faire d'expériences ! Enfin si, peut-être, je peux lancer un dé 100 fois mais cela me prendra beaucoup de temps et le professeur n'appréciera sans doute pas. En plus, je trouverai peut-être une fréquence de 0,63, ce que le professeur n'appréciera sans doute pas non plus.

C'est effectivement un vrai problème. Il faut donner des arguments extra-mathématiques, comme « je ne vois pas de raison a priori pour que le côté pile soit avantagé par rapport au côté face » (ce qui est raisonnable, mais ce qui n'est pas une démonstration au sens mathématique).

Mais en fait, la plupart du temps, on utilise un argument culturel : « il s'agit là d'une situation bien connue, et des experts en lancers de pièce ont paraît-il vérifié que le choix de 0,5 comme probabilité pour pile était un choix adapté à la réalité. Je veux bien les croire sur parole, d'autant plus que je ne vois pas de raison a priori pour que le côté pile soit avantagé par rapport au côté face. »

Ces arguments sont « raisonnables » et donc acceptés sur une copie, mais il faut bien être conscient qu'il ne s'agit pas de démonstrations mathématiques. On les accepte parce qu'on sait bien que de toutes façons la justification du choix d'un modèle ne peut pas être une démonstration mathématique. Par contre, dans le monde mathématique, on ne peut pas utiliser d'arguments du type : « c'est vrai parce qu'un expert l'a dit et que je le crois sur parole ».

Finalement, la modélisation est une activité intermédiaire entre l'observation de la réalité d'une part et le raisonnement mathématique de l'autre. D'une certaine manière, elle combine les difficultés des deux autres activités et nécessite en tout cas une bonne connaissance des deux.

En résumé, le concept de probabilité d'un résultat cherche à idéaliser la notion de fréquence d'obtention de ce résultat dans une expérience réelle ?

Oui, c'est le premier point de vue historique (Pascal, 17ème siècle). Ce point de vue, associé à l'idée d'égalité des chances, conduit à une méthode de calcul qui permet d'obtenir la probabilité d'un événement par un quotient :

$$\text{probabilité d'un événement} = \frac{\text{nombre de cas favorables à l'événement}}{\text{nombre total de cas possibles}}$$

Résoudre des problèmes de probabilités à l'aide de cette formule exige souvent d'utiliser des techniques de *dénombrement* pour évaluer les nombres de cas. Une représentation graphique souvent adaptée pour énumérer les cas est un schéma en forme d'*arbre*.

II Les probabilités conditionnelles

Il existe un deuxième point de vue, plus adapté aux expériences aléatoires uniques (qu'on ne peut pas répéter pour obtenir une fréquence expérimentale), et plutôt fondé sur la

notion d'information.

Voici un exemple :

Je lance un dé. Je ne te montre pas le résultat, mais je te dis que c'est un nombre pair. Quelle est la probabilité que ce soit un nombre supérieur ou égal à 4 ?

Si c'est un nombre pair, c'est 2, 4 ou 6. Dans ce cas, les nombres supérieurs ou égaux à 4 sont 4 et 6. Donc il y a 2 chances sur 3, et la probabilité est $\frac{2}{3}$.

Oui.

Remarque que ce problème est assez différent de ceux qu'on a rencontrés avant. En effet, il pose la question de la probabilité une fois que l'expérience a été faite et non pas avant. De plus, une fois l'expérience faite, il donne une information partielle sur le résultat (il est pair). Ici, les jeux sont faits avant que le joueur parie. Il parie sur ce qui s'est passé et non pas sur ce qui va se passer.

Ta façon de le résoudre a d'ailleurs été différente : tu as répondu comme si on t'avait en fait décrit une autre expérience que celle dont on parlait. Tu as fait comme si le problème était à peu près le suivant : « on dispose de trois cartons numérotés 2, 4 et 6. On en tire un au hasard. Quelle est la probabilité que ce soit un nombre supérieur ou égal à 4 ? ». Avec cet autre problème, tu as alors appliqué la formule du modèle des fréquences (cas favorables / cas possibles) : il y a 2 cas favorables (4,6) sur 3 cas possibles (2, 4, 6), donc la probabilité est $\frac{2}{3}$.

Dans ce type de problème, l'information supplémentaire (le résultat est pair) change la notion de probabilité. Pour en tenir compte, on imagine un autre problème fictif qui incorpore cette information, et on résout cet autre problème.

La probabilité n'apparaît alors plus comme étant intrinsèquement liée à l'expérience initiale (lancer un dé, obtenir un supérieur ou égal à 4), ce qui serait une vision où la probabilité serait *objective, absolue*, liée aux objets et aux événements. Au contraire ici, la probabilité est liée à l'information que possède celui qui parie. Ce serait alors une vision où la probabilité serait *subjective, relative*, elle mesurerait une sorte de quantité de certitude que peut avoir le parieur en fonction des informations dont il dispose.

Effectivement, ces deux points de vue sont différents. mais sont-ils contradictoires ou bien y a-t-il moyen de les réconcilier ?

Il y a effectivement un moyen de les réconcilier, c'est d'introduire la notion de *probabilité conditionnelle* dans le premier point de vue (celui des fréquences).

Il se trouve que le calcul que tu as fait peut s'exprimer par des probabilités obtenues en restant dans le point de vue des fréquences pour l'expérience initiale (lancer un dé).

Dans ce point de vue, obtenir un nombre pair a pour probabilité $\frac{3}{6}$ (3 nombres pairs sur 6 cas possibles en tout).

Ensuite, tenir compte de l'information « le nombre est pair » pour étudier la condition « le nombre est supérieur ou égal à 4 » revient à faire l'intersection de deux ensembles : l'ensemble des nombres pairs $A = \{2, 4, 6\}$ et l'ensemble des nombres supérieurs ou égaux à 4 $B = \{4, 5, 6\}$. l'intersection étant $A \cap B = \{4, 6\}$.

Les probabilités sont alors $P(A) = \frac{3}{6}$ et $P(A \cap B) = \frac{2}{6}$. Ta réponse correspond finalement à l'opération $\frac{P(A \cap B)}{P(A)} = \frac{2}{3}$.

On donne un nom à cette opération, on appelle cela la probabilité conditionnelle de B par rapport à A , ou la probabilité de B sachant A , ce qu'on note $P_A(B)$. Cela traduit le fait qu'on s'intéresse à la probabilité de B non pas dans le contexte initial (ce qui donnerait $P(B) = \frac{3}{6}$) mais dans un contexte où on sait que A est réalisé.

$$P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Cette formule montre comment la connaissance d'une information supplémentaire (A) change la probabilité, qui devient $P_A(B)$ au lieu de $P(B)$. Mais elle permet de la calculer sans avoir besoin d'imaginer une expérience fictive qui tiendrait compte de l'information supplémentaire.

Une probabilité conditionnelle, c'est la probabilité d'une intersection ?

Non. Croire cela est une erreur fréquente. L'intersection intervient dans le calcul, mais le résultat n'est pas $P(A \cap B)$. Autrement dit « B sachant A » n'est pas $A \cap B$. Dans « B sachant A », les rôles de A et B ne sont pas les mêmes. A est certain, c'est une information supplémentaire qu'on a obtenue, alors que B reste hypothétique. Dans $A \cap B$, à la fois A et B sont hypothétiques. Dans « B sachant A », A s'est déjà réalisé, il n'y a pas à se poser de questions à son sujet.

III Conclusion

Voyons si tu as compris : c'est quoi une probabilité ?

Enfinement, c'est simple : une probabilité est une mesure sur une sigma-algèbre d'un univers Ω , telle que la mesure de l'univers soit égale à 1. Pour l'instant, on s'est limité dans les exemples à des sigma-algèbres discrètes sur un univers fini.

En conséquence, $P(A \cup B) = P(A) + P(B)$ si A et B sont des sous-ensembles de l'univers, disjoints.

Et en cas d'équiprobabilité, $P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$

Euh ...

Tu comprendras quand tu seras plus grand ...