

# Théorie des probabilités : une axiomatique, des interprétations

Hugo LAVENANT

12 mars 2015

## Introduction

« On voit que si la géométrie n'est pas une science expérimentale, c'est une science née à propos de l'expérience, que nous avons créé l'espace qu'elle étudie, mais en l'adaptant au monde où nous vivons. »<sup>1</sup>

Ce qu'exprime Henri POINCARÉ à propos de la géométrie pourrait être étendu plus largement aux mathématiques : elles ne décrivent pas fidèlement le réel, elles ne sont pas expérimentales dans le sens où une observation ne pourra jamais *prouver* un théorème ; mais elles s'en inspirent, tout d'abord parce que l'intuition de l'esprit humain s'appuie en grande partie sur ses perceptions directement issues du réel, mais aussi car elles ont été conçues (comme la géométrie) comme une idéalisation du monde sensible.

Donc si l'on veut donner un sens empirique à un résultat abstrait exprimé en langage mathématique, il faut l'interpréter, et cette opération n'est pas simple. Un même résultat mathématique peut avoir *a priori* différentes interprétations. Quand il s'agit de la géométrie, l'interprétation est souvent facile à faire et sans ambiguïtés : un concept comme celui de longueur, qui a un sens mathématique bien précis, se conçoit bien de manière pratique (tout le monde a déjà mesuré des longueurs à l'aide d'un règle) ; c'est pourquoi il est assez difficile de faire percevoir la différence entre longueur mathématique et longueur « empirique ».

Au contraire, la notion d'évènement *probable* est beaucoup plus dure à appréhender, à formuler, ce qui permet de montrer plus facilement la différence qu'il y a entre le concept intuitif de probabilité et son axiomatisation mathématique. Dans ce texte, on essaiera d'expliquer cette différence, et pourquoi il est nécessaire, pour faire des mathématiques, d'évacuer la question de l'interprétation : cette dernière devient une question *métamathématique*.

## 1 Évènement probable et quantification de la probabilité

**Interprétation classique et définition de la mesure de probabilité** La simple consultation d'un dictionnaire donne une idée de la difficulté qu'il y a à définir ce qu'est un évènement probable, même si nous en avons tous plus ou moins une idée intuitive. De manière générale, un évènement *aléatoire* est un évènement dont on ne sait pas avec certitude s'il se produira ou non, c'est le contraire d'un évènement prédictible. L'exemple le plus emblématique est le lancer d'une pièce de monnaie : on ne sait pas à l'avance de quel côté la pièce retombera.

Néanmoins, on sent bien que certains résultats possibles sont moins *probables* que d'autres : le fait que la pièce de monnaie retombe pile sur la tranche semble bien plus improbable comparativement au cas où elle retombe sur un des deux faces. Ainsi, Émile Borel a énoncé la *loi unique du hasard*, comme quoi si un évènement est hautement improbable, alors une intelligence raisonnable doit agir comme si cet évènement n'aura pas lieu.

Ce que tentent les mathématiques, c'est de *quantifier*, de *mesurer* le fait d'être probable. À chaque évènement, on associe un nombre qui mesure la probabilité (c'est-à-dire la fait d'être probable) de cet évènement. Par abus de langage, on appelle ce nombre la probabilité de l'évènement en question, même si on peut signaler que la terminologie mathématique rigoureuse parle précisément de mesure de probabilité de l'évènement. La mesure de probabilité est un élément du modèle

---

1. Henri POINCARÉ, *Science et méthode*.

mathématique et non de l'expérience, c'est-à-dire de la structure abstraite qui cherche à représenter l'expérience soumise à l'aléa. Voici la manière « classique » de relier intuitivement cette mesure à l'expérience elle-même.

On suppose que l'on dispose d'une expérience aléatoire, que l'on peut répéter dans les mêmes conditions, hypothèse importante sur laquelle on reviendra. Cette expérience peut se terminer de différentes façons, et un événement sera un regroupement d'issues possibles de l'expérience. Dans l'exemple du lancer d'une pièce, l'expérience est terminée lorsque la pièce s'immobilise et on peut considérer différents événements : les plus souvent donnés sont « la pièce s'arrête sur pile » ou « la pièce s'arrête sur face » mais on peut en considérer d'autres comme « la pièce passe plus d'une seconde dans les airs » ou encore « la pièce retombe à plus d'un mètre de l'endroit où elle a été lancée ». On répète l'expérience un grand nombre de fois, et pour chaque réalisation de celle-ci, on regarde si l'événement considéré s'est produit. La mesure de probabilité d'un événement est alors la limite du rapport

$$\frac{\text{Nombre d'expériences pour lesquelles l'événement s'est produit}}{\text{Nombre total d'expériences réalisées}}$$

lorsque l'on réalise de plus en plus d'expériences, où l'on discutera du sens à donner à *limite* et *de plus en plus* qui sont ici volontairement flous. Cette mesure est un nombre compris entre 0 et 1, une valeur de 1 signifiant que l'événement arrive presque à coup sûr et une valeur de 0 signifiant qu'il n'arrive presque jamais, et elle est d'autant plus grande que l'événement est probable.

**Une définition pas si claire** L'approche donnée ci-dessus pose finalement plus de questions qu'elle ne donne de réponses. Le concept d'expérience aléatoire n'est pas si évident : c'est une expérience que l'on doit pouvoir répéter *exactement* dans les mêmes conditions autant de fois que l'on veut. Pour poursuivre l'exemple du lancer de la pièce, on ne la lance jamais exactement de la même façon, et c'est d'ailleurs pour ça que le résultat n'est pas toujours le même : une petite variation dans la vitesse initiale avec laquelle on la lance peut changer le résultat final de pile à face. En fait, c'est précisément parce que l'on ne connaît pas les conditions initiales qu'on ne connaît pas le résultat : le calcul des probabilités est une façon de modéliser l'information manquante, l'aléatoire n'est pas intrinsèque mais uniquement dû à notre manque de connaissance<sup>2</sup>. Il existe donc une partie de l'expérience (la façon dont on lance la pièce) qui *doit* varier d'une réalisation à l'autre.

Un autre problème est qu'il faut faire un grand nombre d'expériences, et que la mesure de probabilité d'un événement n'est définie que si on en fait une infinité (si on considère que la mesure de probabilité est une donnée expérimentale que l'on obtient empiriquement). Or, on est souvent limité par le temps ou les moyens et il est complètement impossible d'en réaliser une infinité, ou même un grand nombre (penser à des expérimentations sur les animaux où, pour des raisons éthiques et financières, le nombre de cobayes est restreint). La mesure de probabilité ne peut plus être vue comme obtenue par l'expérience, on aurait envie de la poser comme un *a priori* de la modélisation et serait définie le cas hypothétique d'un nombre infini d'expériences. Mais dans ce cas, on doit *supposer* que les expériences aléatoires possèdent une régularité au lieu de *observer* comme cela peut être le cas dans les jeux de hasard.

## 2 La théorie des probabilités axiomatisée

**Un avant-propos** On va maintenant présenter l'approche mathématique moderne de la théorie des probabilités en montrant comment elle évacue justement la question du sens et de l'interprétation qu'elle donne aux mesures de probabilité. L'axiomatique qui va suivre a été proposée par Andreï KOLMOGOROV<sup>3</sup> dans les années 1930, même si auparavant de nombreux mathématiciens avaient déjà travaillé sur le calcul des probabilités : on peut penser à Blaise PASCAL ou Pierre Simon DE LAPLACE. Mais les axiomes, les hypothèses sous-tendant leurs différents travaux différaient, justement parce qu'ils étaient liés à des questions d'interprétation du concept de mesure de la probabilité, qui, comme on l'a vu plus haut, n'est pas si facile à définir.

2. Certains pourraient argumenter que dans le cas de la mécanique quantique l'aléatoire est réellement intrinsèque et que l'on reproduit l'expérience *exactement* de la même façon. Même si cela est exact, il faut bien reconnaître que le calcul des probabilités ne s'utilise pas qu'en mécanique quantique mais aussi dans des cas comme les jeux de hasard où l'aléatoire modélise l'information manquante.

3. Andreï KOLMOGOROV (1903-1987), mathématicien russe.

Une mise en garde s'impose : les termes employés par la suite comme *univers*, *événements*, *mesure de probabilité*, renvoient à des idées intuitives, primitives que l'on a à propos des probabilités mais sont d'un point de vue mathématique une simple terminologie qui ne dit *rien* à propos des objets mathématiques qu'ils désignent. C'est-à-dire qu'on pourrait remplacer dans la suite du texte ces mots par tables, chaises, fauteuils etc. sans que la rigueur mathématique s'en trouve changée. En fait ce qui importe ce n'est pas ce que *sont* les objets mathématiques mais la façon dont ils sont en *relation* :

« on peut se rendre compte que le rôle des entités prime sur leur *nature* et que l'essence est contemporaine de la relation. »<sup>4</sup>

Dans la présentation de l'axiomatique qui suit, on utilisera comme illustration afin de faciliter la compréhension l'expérience correspondant au lancer d'un dé à six faces, et l'on ne s'intéresse qu'au numéro de la face qui est au-dessus lorsque le dé s'immobilise.

**La théorie proprement dite** Une expérience aléatoire est modélisée par l'ensemble de ses issues, c'est-à-dire des manières dont elle peut se terminer, cet ensemble est appelé l'*univers*. Dans notre exemple, l'ensemble des issues possibles est codé par  $U, D, T, Q, C, S$ , la lettre désignant la première lettre du chiffre tiré. L'univers est noté  $\{U, D, T, Q, C, S\}$ .

On appelle *événement* tout regroupement d'issues, c'est-à-dire toute partie de l'*univers*. Par exemple, l'événement « le chiffre sorti est pair » est codé par le sous-ensemble  $\{D, Q, S\}$  et l'événement « le chiffre sorti est multiple de trois » est codé par  $\{T, S\}$ . Il existe deux événements particuliers : l'*événement certain* correspondant à l'univers tout entier (dans notre exemple  $\{U, D, T, Q, C, S\}$ ) et l'*événement impossible*, quand aucune issue n'est réalisée (il est codé par l'ensemble vide  $\emptyset$ ).

Enfin, une *mesure de probabilité* (notée  $\mathbb{P}$ ) est une fonction qui à tout événement associe un nombre entre 0 et 1 et qui satisfait les trois conditions suivantes :

1. La mesure de probabilité de l'événement certain est 1.
2. La mesure de probabilité de l'événement impossible est 0.
3. Si deux événements sont disjoints (c'est-à-dire qu'ils ne peuvent se réaliser simultanément), alors la mesure de probabilité qu'un des deux événements au moins se réalise est égale à la somme de la mesure de probabilité des deux événements ; ce qui se note formellement

si deux événements  $A$  et  $B$  sont disjoints alors  $\mathbb{P}(A \text{ ou } B) = \mathbb{P}(A) + \mathbb{P}(B)$ .

La première condition stipule que l'événement certain (contenant toutes les issues possibles) est précisément certain, c'est-à-dire associé à une mesure de probabilité 1 ; tandis que la deuxième dit que l'événement impossible (qui ne pourra jamais se réaliser) est associé à une mesure de probabilité 0. La troisième condition, la plus cruciale, permet simplement d'écrire des choses comme la suivante : l'événement « le chiffre sorti est un multiple de trois » est la réunion de deux événements disjoints, à savoir « le chiffre sorti est trois » et « le chiffre sorti est six » ; en conséquence sa mesure de probabilité est simplement la somme de la mesure de probabilité que le chiffre sorti soit trois et que le chiffre sorti soit six.

Et ces trois règles suffisent. C'est-à-dire qu'hormis quelques conditions techniques dans le cas où l'univers comporte une infinité d'issues<sup>5</sup>, un mathématicien appellera expérience aléatoire la donnée de l'univers, des événements et d'une mesure de probabilité satisfaisant aux règles énoncées ci-dessus. Il n'y a absolument pas besoin que ces objets mathématiques soient reliés de près ou de loin à une expérience « réelle », ce qui compte est uniquement la relation que ces objets entretiennent les uns envers les autres.

**Un sens absent** Illustrons comment on s'abstrait du sens par la notion d'*indépendance*. Intuitivement, deux événements sont indépendants s'il n'y a pas de lien entre eux, si la réalisation du premier n'influe pas sur le second. Par exemple, si on lance deux dés l'un après l'autre, on a envie de croire que le résultat du premier et du second dé sont indépendants. Et en ce qui concerne la

4. Gaston BACHELARD, *Le nouvel esprit scientifique*, à propos de la « l'idée philosophique générale de la nouvelle liberté mathématique ».

5. Dans ce cas, les événements seront des regroupements d'issues particuliers et la troisième règle est modifiée de façon à pouvoir considérer une infinité d'événement deux à deux disjoints, mais en donner un énoncé rigoureux demande des outils mathématiques plus sophistiqués.

mesure de probabilité, cela signifie par exemple que la probabilité que la premier dé s'arrête sur le chiffre un et le deuxième sur le chiffre trois est égal au produit des probabilités précédentes ; du moins c'est ce qu'on espère. Les mathématiciens prennent le problème à l'envers : *par définition*, deux évènements mathématiques sont indépendants si la mesure de probabilité que les deux se réalisent est égal au produit de la mesure de probabilité de chacun des deux évènements ; ou pour l'écrire formellement

deux évènements  $A$  et  $B$  sont indépendants si  $\mathbb{P}(A \text{ et } B) = \mathbb{P}(A) \times \mathbb{P}(B)$ ;

sans que cela ne signifie quelque chose sur la manière dont les deux évènements pourraient être liés dans une expérience réelle.

Les questions auxquelles le mathématicien répondra n'auront pas non plus de sens intuitif, elles prennent souvent la forme : connaissant la mesure de probabilité de certains évènements, peut-on calculer la mesure de probabilité d'autres évènements ? Par exemple, si la mesure de probabilité d'obtenir le chiffre trois est  $0,1$  et celle d'obtenir le chiffre six est  $0,3$  alors la mesure de probabilité d'obtenir un multiple de trois est  $0,1 + 0,3 = 0,4$ . Cette déduction est purement mathématique et n'apprend *a priori* rien sur une expérience réelle. En fait, le mathématicien ne quitte jamais le monde mathématique, ses déductions logiques se font *toujours* en utilisant les règles de l'axiomatique qu'il s'est donné et *jamais* avec des raisonnements intuitifs sur des expériences aléatoires réelles<sup>6</sup>.

**Une axiomatique féconde** Ce jeu d'axiomes est utilisé par tout les mathématiciens qui travaillent en théorie des probabilités, et quel que soit le type de phénomènes qu'ils veulent modéliser. Il a permis la reconnaissance de la théorie des probabilités comme une théorie mathématique à part entière<sup>7</sup> et a considérablement augmenté le nombre de problèmes auxquels elle peut s'attaquer. Avec des bases plus rigoureuses, on peut formuler des questions mathématiques plus précises et y répondre, on s'affranchit des chaînes qui liaient les mathématiques à l'expérience, la logique et l'axiomatique deviennent les seules contraintes.

Et l'on voit que la question de l'interprétation est complètement évacuée : on *ne* cherche *pas* à savoir ce que signifie la mesure de probabilité d'un évènement, on sait seulement comment sont *liées* les mesures de probabilité des différents évènements entre elles. Et comme on va le voir, sur cette axiomatique peuvent se greffer différentes interprétations, toute cohérentes avec elle. La science mathématique, en s'intéressant à ses propres questions, refuse de donner un sens unique à ses réponses.

### 3 Différentes interprétations de la théorie des probabilités

**L'application à l'expérience : la modélisation** Le premier avantage de l'approche axiomatique est qu'elle permet de séparer clairement le travail mathématique de celui de modélisation. Les résultats démontrés par un mathématicien sont *prouvés* indépendamment de ce qu'ils signifient, n'importe qui peut les vérifier en suivant le raisonnement logique sans discuter des questions d'interprétation.

Lorsqu'on se trouve en face d'un phénomène que l'on veut *modéliser* à l'aide de la théorie des probabilités, on a à notre disposition plusieurs manières de le faire : la théorie des probabilités elle-même ne peut pas nous dire quel modèle est le meilleur. Un modèle n'est ni vrai, ni faux, il est plus ou moins adapté. Cela donne beaucoup plus de liberté, et ouvre de nouvelles façons de penser. Les mathématiques ne sont qu'une manière d'appréhender le réel, elles n'ont pas la prétention de dicter la *vérité* à son propos.

**Des formes d'aléatoire hétérogènes entre elles** Pour montrer à quel point le sens ne joue pas, on présente ici une petite liste (absolument pas exhaustive) de domaines différents où la théorie mathématique des probabilités est utilisée comme outil de modélisation, on pourra se rendre compte à quel point la notion d'aléatoire varie d'un exemple à l'autre.

6. En réalité, il s'inspire de l'idée intuitive qu'il a de l'expérience, mais il la transforme en arguments de nature mathématique.

7. auparavant, considérée comme non rigoureuse, les mathématiciens ne lui accordaient pas le statut de théorie mathématique.

- Les jeux de hasard : il s'agit de l'exemple discuté dans la première section, où l'aléatoire est vu comme un manque d'information sur une partie qu'on ne contrôle pas (comme le lancer d'un dé). On peut réaliser (de façon opératoire) la même expérience un grand nombre de fois, la mesure de probabilité d'un événement correspond alors à la limite de la fréquence des réalisations de celui-ci.
- La prévision météorologique : par exemple quand on dit qu'il y a 20% de chance qu'il pleuve demain. Même si l'aléatoire est proche de celui des jeux de hasard, on ne peut réellement reproduire la même expérience (les grandeurs physiques indiquées par les capteurs des différentes stations météorologiques ne seront jamais tous identiques à un même instant dans le passé), cela n'a donc pas de sens de parler de la mesure de probabilité comme d'une limite de fréquences quand on reproduit une même expérience.
- La mécanique quantique : contrairement aux jeux de hasard, l'aléatoire est considéré comme intrinsèque, il n'est pas dû à un manque d'information, il n'y a réellement *aucun* moyen de connaître le résultat à l'avance.
- La modélisation du cours de la bourse : l'aléatoire est dû au fait que l'on connaît pas le comportement des autres acteurs du marché. L'irrationalité du comportement humain, les émotions etc. sont modélisés par une mesure de probabilité attribuées aux choix possibles des différents acteurs (donc l'univers mathématique sera l'ensemble des choix possibles, un événement mathématique sera un regroupement de différents choix, et la mesure de probabilité d'un choix quantifie la tendance qu'a un acteur à faire ce choix). On sent tout de suite le manque de réalisme du modèle<sup>8</sup>, ou du moins que cela n'a rien à voir avec un lancer de dé.

Et, au risque de se répéter, tous ces aléatoires sont modélisés à travers le même formalisme mathématique.

**Une interprétation non orthodoxe : la statistique bayésienne** Discutons maintenant de l'approche bayésienne de la statistique, où l'on attribue une mesure de probabilité sans qu'il n'y ait nécessairement d'aléatoire dans les phénomènes. En effet, on attribue une mesure de probabilité à des hypothèses scientifiques qui mesurent le degré de confiance *subjectif* que l'on accorde à ces hypothèses en fonction de ce que l'on sait à l'instant présent. C'est-à-dire que compte tenu des théories, des observations dont on dispose à l'heure actuelle, on se demande s'il est raisonnable de croire à une hypothèse ou non. Et l'on quantifie à quel point il est raisonnable d'y croire à l'aide d'un nombre qui sera la mesure de probabilité de cette hypothèse.

Le calcul des probabilités intervient quand il s'agit de mettre à jour la mesure de probabilité d'une hypothèse lorsqu'on a fait une nouvelle observation. Et dans ce cas, ce qui joue le rôle d'un événement mathématique est une *hypothèse* (c'est-à-dire que l'on calcule la mesure de probabilité d'une hypothèse), les règles satisfaites par la mesure de probabilité sont les mêmes, on effectue des calculs purement formels pour mettre à jour notre degré de confiance en fonction des observations. Et l'on voit qu'on ne peut absolument pas relier cette mesure de probabilité à l'interprétation donnée dans la première section : il n'est pas question d'une expérience qu'on referait un grand nombre de fois puisque l'hypothèse est soit vraie soit fautive, il n'y a objectivement pas d'aléatoire, l'incertitude provient de notre connaissance subjective.

Donnons un exemple afin de clarifier le paragraphe ci-dessus, d'après un calcul fait originellement par LAPLACE. Depuis que je suis né, j'ai vu le soleil se lever chaque matin. Au vu de cette information, est-il raisonnable de croire qu'il va se lever le jour suivant ? (au jour où j'écris ce texte, la question est : le soleil va-t-il se lever le 8 mars 2015 ?)<sup>9</sup>. Cette expérience n'est pas reproductible car il n'y a qu'un seul 8 mars 2015. Le 9 mars 2015 sera un jour différent puisque je serais en possession d'une information supplémentaire, à savoir si le soleil s'est levé ou non le 8 mars. Pour les curieux, compte tenu de mon âge je trouve que la mesure de probabilité de l'hypothèse « le soleil se lève le 8 mars 2015 » est approximativement 0,999875.

8. même si un modèle n'a pas besoin d'être réaliste pour être prédictif et donc utile.

9. Ici, je suppose que je fais abstraction de toutes les connaissances que je peux avoir sur le mouvement des astres, et que l'observation comme quoi le soleil s'est levé tous les matins depuis ma naissance est la seule information dont je dispose. La probabilité que je calcule est donc moins élevée que si de tenais compte d'autres théories (mécanique classique, optique etc.).

## 4 De la difficulté de l'apprentissage des probabilités

La théorie des probabilités a cette particularité qu'il est particulièrement dur d'en donner une interprétation simple et intuitive, ce qui fait qu'il a ralenti son progrès puisqu'on ne savait pas précisément quel objet on manipulait. Au contraire, les objets de la géométrie ont une interprétation intuitive immédiate, ce qui fait que même si une axiomatisation propre et rigoureuse n'est pas venue tout de suite, on a pu accomplir de grands progrès en géométrie puisque tout le monde s'accordait sur les manipulations que l'on pouvait effectuer sur les objets géométriques.

L'enseignement au lycée, ou maintenant en classes préparatoires, tient à donner un aspect formalisé et rigoureux de la théorie. Mais il ne discute pas assez, à mon sens, que l'application à des expériences réelles est une question de *modélisation*. On laisse donc l'élève seul face à un exercice où une « mise en contexte », décrivant une expérience aléatoire, lui est présenté et il doit faire lui-même le travail de modélisation. Voici un exemple d'énoncé qui pourrait faire un exercice de niveau terminale en rajoutant des questions intermédiaires :

*Une personne possède  $n$  clés et se présente devant une porte. Une seule des  $n$  clés permet d'ouvrir la porte.*

- Il essaye les clés les unes après les autres en éliminant à chaque essai la clé si elle n'a pas convenue. Déterminer le nombre moyen d'essais nécessaires pour ouvrir la porte.*
- Il essaye les clés les unes après les autres sans éliminer celles qui ne conviennent pas. Déterminer le nombre moyen d'essais nécessaires pour ouvrir la porte.*

Ici, identifier un univers, des événements, une mesure de probabilité est en réalité assez compliqué. De plus, ils n'ont pas besoin d'être explicités pour résoudre le problème, une correction standard n'y fera pas référence, préférant des raisonnements mi-intuitifs et mi-mathématiques.

On peut remarquer au contraire qu'à mesure que l'on avance dans les études les énoncés des problèmes se font de plus en plus abstraits et peuvent plus ou moins se mettre sous la forme

*Soient un univers, des événements et une mesure de probabilité satisfaisant certaines relations supplémentaires. Déterminer la probabilité de tel événement.*

avec éventuellement une introduction au début de l'exercice pour savoir pourquoi on considère cette mesure de probabilité en particulier, mais sans que cela ne soit vraiment utile pour répondre aux questions posées.

Le principal problème soulevé par cette approche est que les élèves ont du coup beaucoup de mal à faire le lien entre leur cours très formel et les exercices plus terre-à-terre. De plus, la modélisation se faisant parfois de manière subtile, il n'est pas rare qu'ils fassent des erreurs d'interprétation (par rapport à ce qui est attendu par le concepteur de l'exercice) sans que ces dernières soient réellement mathématiques, on devrait plutôt les qualifier de *métamathématiques*.

Il n'est pas question de remettre en question le fait de d'enseigner la théorie des probabilités, au contraire elle entre en compte dans de nombreux processus de modélisations et est indispensable pour comprendre les statistiques qui sont une rare partie des mathématiques dont n'importe qui peut réellement entendre parler tous les jours. Mais ce serait sûrement mieux si on différenciait réellement mathématique et modélisation, si on insistait sur le fait que la théorie des probabilités a pris son envol mathématique à partir du moment où elle a développé une axiomatique et s'est coupée des questions d'interprétation ; fait dont, je l'espère, ce texte vous aura convaincu.