L'ENSEIGNEMENT DU CALCUL DES PROBABILITÉS DANS LE SECOND DEGRÉ

Perspectives historiques, épistémologiques et didactiques

Michel HENRY
Irem de Besançon

"Nul n'entre ici s'il n'est géomètre"

Cette apostrophe, gravée par Platon au fronton de l'Académie, attribue aux mathématiques un rôle de fondement de toute connaissance et un statut éducatif essentiel. A la base de la pensée platonicienne (*La République*, Livre VI, 380 av. J.C. [1]) (1), il y a la distinction entre le monde des objets sensibles, imparfaits, complexes et changeants, et le monde de leurs modèles éternels, à savoir le monde des Idées, parfaites et immuables. Ce monde des Idées est seul intelligible et, dès lors le raisonnement déductif peut s'y développer en toute sécurité.

L'enseignement de la géométrie a fait l'objet de nombreuses études didactiques mettant en évidence cet *enchaînement de contrats* qui, lors de leurs premiers apprentissages, consiste à placer les élèves dans le domaine du sensible, à les entraîner à l'observation géométrique et à la manipulation des instruments de dessin, puis, par une sorte de rupture impliquant un effort intellectuel que certains acceptent difficilement, à leur demander de passer dans le domaine des idées, pour s'engager dans des démonstrations donnant aux figures un nouveau statut d'abstractions régies par des hypothèses.

L'enseignement des probabilités fait l'objet de réflexions analogues. Les récents travaux didactiques (citons Sylvette Maury [32], Jacques Bordier [33], Moncef Zaki [34] et André Totohasina [35]) étudient les conceptions spontanées qui se dressent en obstacles à la compréhension des situations probabilistes.

En probabilités, ces conceptions spontanées s'expriment dans un contexte où l'irrationnel n'est pas absent, notamment lorsque

(1) Voir sur ce point le livre de Jean Dhombres et de l'IREM de Nantes, Nombres mesure et continu, Cedic 1978 [21]

des problèmes de prédiction sont posés. Leur étude est donc particulièrement importante dans ce domaine où s'applique encore plus que dans d'autres la remarque de Gaston Bachelard [19, p. 23] : *"Dans la formation de l'esprit scientifique, le premier obstacle, c'est l'expérience première"*.

La confusion encore entretenue dans l'enseignement entre les propriétés du hasard, les combinaisons des événements aléatoires, et le modèle mathématique construit autour du concept de probabilité, provoque par exemple :

— le refus de pouvoir dire quoi que ce soit de fiable sur l'avenir,

— de nombreuses questions sur le caractère objectif ou subjectif de la probabilité,

— la confusion entre probabilité et possibilité d'observer un événement,

— l'identification formelle entre probabilité et fréquence,

— le sentiment d'être en présence d'un "cercle vicieux" lorsque l'on démontre la loi des grands nombres après avoir introduit le concept de probabilité par son lien avec la fréquence d'un événement basé sur une appréhension spontanée de cette même loi, comme le propose le programme de première.

— l'incompréhension de paradoxes comme celui des trois bancs présenté ensuite, ou celui de Joseph Bertrand (2) [11],

— l'attribution d'une probabilité à des événements non aléatoires,

— ou la recherche dans la nature des lois de variables aléatoires classiques idéalisées.

Un grand pas en avant est franchi dans cet apprentissage quand les élèves savent distinguer la description du réel, avec ses objets imparfaits, complexes et changeants, du modèle probabiliste avec ses concepts, ses définitions, ses hypothèses et ses théorèmes.

La difficulté d'un problème de probabilités réside alors dans la construction du modèle adéquat à partir des données de l'observation — ou de l'habillage pseudo-concret que l'enseignant a donné à ce problème.

Une fois ce modèle établi, le reste n'est que mathématiques. La reconnaissance des bons outils probabilistes en présence de situations aléatoires déterminées, conduisant à des résultats pertinents et opératoires en pratique, relève de l'habileté du probabiliste et de sa pratique professionnelle. Cette dernière requiert quelques temps d'assimilation.

Une autre difficulté, mais de moindre ampleur, repérée par les enseignants, est la complexité de certains problèmes de logique combinatoire. Cette difficulté didactique devenait principale quand l'hypothèse primitive du modèle probabiliste était l'équiprobabilité des événements élémentaires (et nous verrons dans la suite que cette hypothèse a été historiquement nécessaire du 17e au 19e siècle) réduisant le calcul des probabilités enseigné dans le second degré à des situations plus ou moins simples de dénombrements.

Dans le second degré, les programmes actuels ont élargi cette conception en laissant ouvertes

(2) Paradoxe du choix au hasard d'une corde d'un cercle : quelle est la probabilité qu'elle soit plus longue que le côté du triangle équilatéral inscrit dans ce cercle ? La loi du choix au hasard n'étant pas précisée, on peut obtenir diverses valeurs pour cette probabilité. Ce paradoxe est présenté dans l'article "Paradoxes et lois de probabilité" *in Repères-IREM* n° 13, Janvier 1994.

ces probabilités élémentaires aux cas d'événements "*non également possibles* ", selon les termes de Laplace [9, p. 38]. L'importance des dénombrements, source d'échecs en probabilités au niveau du second degré, est ainsi relativisée.

Au niveau de l'enseignement supérieur, notamment en Licence, le choix actuel est de se placer d'emblée dans un modèle mathématique supérieur pour y établir des résultats puissants. La plupart des problèmes évitent ainsi les obstacles de la construction de ce modèle, les contrôles d'adéquation et les affres des dénombrements. Il ne reste comme difficulté que l'assimilation rapide de notions nouvelles et la mise en œuvre de techniques mathématiques de haut niveau que l'on suppose maîtrisées.

I — L'enseignement des probabilités au Lycée

1. Les objectifs des programmes rénovés de 1991

a) L'approche fréquentiste et la progression :

Dans la continuité du programme de seconde, ceux de première et de terminale proposent de s'appuyer sur la description d'une série statistique pour introduire la notion de probabilité.

Mais, et ceci est essentiel, une probabilité ne peut avoir de sens que si elle concerne un événement associé à une expérience aléatoire.

En l'absence d'expérience aléatoire, pas de probabilité.

Le concept de probabilité, et sa construction, suppose donc acquis ceux d'expérience aléatoire et d'événements. De plus, l'approche est résolument fréquentiste, comme le montre l'introduction du programme de première :

"L'objectif est d'entraîner les élèves à ***décrire quelques expériences aléatoires simples*** *et à calculer des probabilités. On évitera tout développement théorique. Pour introduire la notion de probabilité,* ***on s'appuiera sur l'étude de séries statistiques*** *obtenues par répétition d'une expérience aléatoire, en soulignant les propriétés des fréquences* ***et la relative stabilité de la fréquence*** *d'un événement donné lorsque cette expérience est répétée un grand nombre de fois."*

Comme les techniques de dénombrement ne figurent qu'au programme de terminale, le caractère fréquentiste de cet enseignement est confirmé. Cela ne va cependant pas sans quelques ambiguïtés épistémologiques résiduelles ni sans difficultés didactiques, comme nous le verrons ensuite.

Ainsi, dans leur progression, les programmes des seconde, première et terminale abordent les points suivants :

classe de seconde :

— étude descriptive des séries statistiques, diagrammes représentatifs, paramètres de position et de dispersion, séries classées, mode, effectifs et fréquences.

classes de première :

— descriptions d'expériences aléatoires simples,
— études expérimentales de la stabilisation des fréquences lors de la répétition d'une même expérience aléatoire,
— notion de probabilité dans une conception objective et fréquentiste,
— modèle mathématique du calcul des probabilités, vocabulaire des ensembles et propriétés

élémentaires des probabilités,

— éléments de calcul des probabilités d'événements combinés,

— cas particulier de l'équiprobabilité, classes de terminales :

— notion de probabilités conditionnelles,

— notion de variables aléatoires et de lois de probabilités,

— éléments de combinatoire, applications à des calculs de probabilités.

b) Critique didactique :

Du point de vue de la transposition didactique, la probabilité d'un événement A apparaît comme valeur théorique "limite" des fréquences des réalisations de A lors de la répétition de la même expérience aléatoire un grand nombre de fois.

Cet objectif fait qualifier de fréquentiste cette approche de la notion de probabilité. Il est ambitieux et d'une certaine manière ambigu.

Il est ambitieux dans la mesure où il propose d'atteindre directement toute l'étendue du concept, sans les étapes intermédiaires indiquées dans la suite de cet article.

Il cumule des difficultés didactiques en associant la probabilité à une notion de "limite", plus complexe qu'en analyse (puisqu'on parle de "stabilisation relative"), alors que se pose encore le problème de l'existence même de cette limite dans la tête des élèves. On pourrait faire le parallèle avec l'introduction de la notion de réel par limite de suites décimales illimitées. On sait que cette ambition s'est révélée inaccessible dans le second degré.

La non-maîtrise du vocabulaire des ensembles ajoute à la difficulté de compréhension par les élèves des situations probabilisées. Ce vocabulaire n'est en général pas introduit avant le chapitre sur les probabilités.

Enfin, et c'est la source d'obstacles principaux, l'introduction de la notion de probabilité se heurte à des conceptions préconstruites, erronées ou même contradictoires qui commencent à être bien étudiées dans la littérature didactique (Voir par exemple [32, 34 ou 35]).

Le programme conserve une part d'ambiguïté en ne fixant pas comme objectif explicite la distinction entre les situations aléatoires réelles et le modèle mathématique censé les décrire. Pour les enseignants non prévenus, ce programme tend à confondre fréquence et probabilité. Plus précisément, il confond deux domaines au sens platonicien : le domaine du sensible dans lequel les descriptions statistiques prendront toute leur place, et le domaine des Idées où la théorie peut se développer sur la base d'hypothèses abstraites.

En ne séparant pas suffisamment la probabilité mathématique de sa signification concrète, l'outil introduit semble alors mis en cause par les choix épistémologiques voir philosophiques des sujets, d'une part sur la place du hasard dans la nature (degré de notre ignorance ou phénomène objectif ?) d'autre part sur la notion intuitive de probabilité pénétrée de tendances subjectives (les "chances" que j'ai d'obtenir tel ou tel événement).

Enfin, malgré son introduction fréquentiste du concept de probabilité, les élèves ont spontanément recours au calcul a priori d'une probabilité, lorsque la situation se prête simplement à une approche de "géométrie du hasard" (selon les termes de Pascal) lorsque son estimation a posteriori par une étude statistique des fréquences se révèle trop lourde. La question du sens à donner au concept de probabilité reste posée.

Dans le cas particulier où l'hypothèse d'équiprobabilité sur les événements élémentaires est licite, cette probabilité coïncide avec le rapport du nombre d'événements réalisant A (cas favorables) au nombre des événements élémentaires (cas possibles).

Le problème didactique est alors de faire comprendre clairement cette identification.

C'est notamment le cas dans les situations de tirage "au hasard" d'un élément d'une population. L'hypothèse d'équiprobabilité sur cette population identifie alors la **probabilité** d'obtenir un élément de qualité A donnée à la **proportion** des éléments ayant cette qualité dans la population. Cette identification là ne semble pas poser de problème aux élèves, et les expérimentations montrent clairement qu'elle est spontanée ou préconstruite. Elle repose en effet sur la conception que *dans une situation aléatoire parfaitement symétrique donnant lieu à n issues, chaque issue a "une chance sur n"*. L'obstacle se révèle quand les élèves adoptent la même attitude lorsqu'il n'y a pas de symétries et seulement n éventualités.

2. Introduction du concept de probabilité

a) du pourcentage à la probabilité

Les remarques précédentes ont semble-t-il inspiré les rédacteurs de programmes, notamment celui qui fut préparé pour la série L. Quel que soit le sort qui lui a été réservé, ce texte reste intéressant par la clarté de ses objectifs, il avait d'ailleurs reçu un avis très favorable de la commission inter-IREM "statistiques et probabilités". Voici son introduction :

"Il s'agit, comme dans les autres programmes, d'aborder la notion de probabilité à partir de la fréquence, mais on a choisi dans cette série d'affiner l'explication du processus de ***modélisation****. L'objet de cette partie de la formation est donc de faire découvrir, en s'appuyant sur l'expérimentation numérique, quelques notions qualitatives et quantitatives liées à la modélisation mathématique des phénomènes aléatoires.*

L'objectif est d'abord de dégager, de façon ***qualitative*** *et éventuellement critique, quelques notions probabilistes en répertoriant et en analysant des situations dans lesquelles sont mis en œuvre ou observés des phénomènes aléatoires ou d'apparence aléatoire.*

On abordera ensuite une analyse plus ***quantitative*** *permettant de dégager à partir de l'étude des fréquences et en relation avec les travaux effectués en statistique, la notion de probabilité. On fera également le lien avec le calcul des pourcentages".*

Le suivi de l'enseignement des probabilités dans des classes de premières [3] et terminales me conduisent à quelques remarques didactiques, qui devraient être étayées par les recherches en cours.

La critique précédente suggère d'aborder la notion de probabilité par cette identification restrictive à la proportion, le hasard étant matérialisé par le tirage aléatoire d'un élément dans la population étudiée.

Cette identification est, semble-t-il, l'argument le plus fort pour détruire la conception erronée très répandue chez les élèves, qui tendent à confondre le nombre des éventualités qui peuvent sortir d'une expérience aléatoire et les "chances" que l'on a de les observer.

(3) Pour une analyse détaillée de ce programme et la description d'une séquence d'enseignement, on pourra se reporter au bulletin de l'IREM de Besançon n° 45 (Juin 91) ou à l'article paru dans *Repères-IREM* n° 6 (Janvier 92)

La probabilité d'un événement, issue dans un premier temps de la proportion des objets de caractère A dans une population où l'on a effectué un tirage au hasard, est ensuite associée au dénombrement des "cas favorables" réalisant A dans une expérience aléatoire, où l'hypothèse d'équiprobabilité des événements élémentaires ne fait pas de doute ; la probabilité de A apparaît alors comme la proportion des événements élémentaires réalisant A parmi toutes les issues possibles de l'expérience, constituant la population statistique étudiée.

Cette extension s'appuie sur la conception spontanée que *la réalisation de A est d'autant plus probable que la proportion des éléments de caractère A est grande dans la population.*

Le concept de probabilité est ainsi élargi à des expériences aléatoires qui ne se réduisent pas directement à des tirages "au hasard" dans une population. Bien que dans la plupart des cas (où il y a équiprobabilité quelque part) on puisse introduire un modèle d'urne, cela ne semble pas toujours pertinent sur le plan didactique, car on risque alors un glissement analogique et un effet distracteur vis à vis de la compréhension de la situation.

Notons que l'hypothèse d'équiprobabilité, lorsque l'expérience aléatoire n'est pas suffisamment décrite, peut être la source d'ambiguïtés et de petits paradoxes élémentaires, comme on peut le voir dans l'exemple suivant, communiqué par P. Perrin de l'IUFM de Reims :

"Dans un square se trouvent trois bancs de deux places chacun. Deux personnes arrivent et s'assoient au hasard. Quelle est la probabilité qu'elles s'assoient côte à côte ?"

Si chaque personne tire un banc au hasard, la probabilité est 1/3,

si chaque personne tire une place au hasard, la probabilité est 1/5. (On suppose que lorsque la même place est tirée par les deux personnes, elles refont leur choix).

b) vers la conception fréquentiste

Puis, dans une troisième étape, les élèves peuvent être amenés à constater qu'en dehors des situations modélisées par des tirages d'urnes ou des jeux de hasard, il y a peu de situations naturelles où l'équiprobabilité a un sens (phénomènes météorologiques, propagation des maladies, choix stratégiques d'un joueur, sont les exemples développés par Jacques Bernoulli dans "*Ars Conjectandi*" [5, p. 42]).

La nécessité d'élargir le concept de probabilité se présente avec force et des activités comme celle du lancer de punaises (à la mode dans les classes cette année) ou comme le comptage des files d'attente nécessitent l'approche fréquentiste pour estimer (*i.e.* donner une valeur numérique approchée à) la probabilité qui reste un *nombre théorique attaché objectivement à l'expérience en cours et à l'événement attendu.*

Ce nombre étant préconçu grâce aux exemples simples issus de tirages au hasard, les élèves acceptent alors plus facilement l'idée de la stabilisation des fréquences observées vers des valeurs de plus en plus proches de cette probabilité théorique.

Ils peuvent alors adopter vis à vis de cette probabilité une attitude familière en physique, considérant un résultat de mesure comme une valeur approximative (et entachée d'erreur) d'une mesure théorique de la grandeur étudiée.

Cette généralisation de la notion de probabilité repose sur la reconnaissance du fait que *lorsqu'on répète une expérience aléatoire un grand* *nombre de fois, la proportion (ou fréquence) observée des réalisations d'un événement A est voisine de sa probabilité a priori.*

Ce fait est facilement anticipé par les élèves dans les situations simples comme pile ou face ; ils l'acceptent comme moyen de contrôle d'un raisonnement a priori dans des cas plus complexes comme le lancer de deux pièces identiques ; ils comprennent son extension aux situations qui ne se ramènent pas à un système de cas également probables, comme par exemple l'aide au diagnostic en présence de symptômes indicateurs de plusieurs pathologies.

c) le lien probabilité-fréquence

Parmi les cours s'inspirant de l'approche fréquentiste, il faut citer la "*théorie des probabilités*" d'Hélène Ventsel publié en 1973 aux éditions Mir de Moscou [30]. Dans sa présentation, Hélène Ventsel étudie longuement la notion de fréquence pour estimer les probabilités, elle souligne [30, p. 23] : "*toutes ces méthodes* [d'estimation] *ont leurs racines dans l'expérience...*"

La clarté de son exposé mérite une lecture un peu détaillée [30, p. 24] :

"Le lien existant entre la fréquence d'un événement et sa probabilité est très profond. [...] En effet, lorsqu'on estime la probabilité d'un événement, on se base forcément sur la fréquence des événements analogues dans la pratique. En caractérisant la probabilité d'un événement par un nombre quelconque on ne peut attribuer à ce nombre une autre valeur réelle et un autre signification que la fréquence relative d'apparition de l'événement en question dans un grand nombre d'expériences.

L'estimation numérique de l'éventualité d'un événement par la probabilité n'a de sens pratique que ***parce que les événements plus probables se produisent en moyenne plus souvent que les moins probables.*** *Et si l****a pratique montre que****, au fur et à mesure de l'augmentation du nombre d'expériences, l****a fréquence de l'événement a tendance à se stabiliser,*** *s'approchant rigoureusement d'un nombre constant, il est naturel d'admettre que ce nombre est la probabilité de cet événement.*

Il est clair que l'on peut vérifier cette hypothèse seulement pour des événements dont les probabilités sont susceptibles d'être calculées directement, c'est à dire pour les événements se réduisant à un système de cas, car ce n'est qu'alors qu'il existe une méthode permettant de calculer la probabilité mathématique.

De nombreuses expériences réalisées depuis la naissance du calcul des probabilités confirment effectivement cette hypothèse. ***Il est tout naturel d'admettre que pour les événements ne se réduisant pas à un système de cas, la même loi reste vraie*** *et que le nombre constant vers lequel tend la fréquence d'un événement, lorsque le nombre d'expériences augmente, n'est rien d'autre que la probabilité de l'événement. On peut alors prendre la fréquence d'un événement, calculée pour un nombre suffisamment grand d'expériences, pour la valeur approchée de la probabilité. C'est ce qu'on fait en pratique, en déterminant à partir de l'expérience les probabilités des événements ne se réduisant pas à un système de cas".*

Peut-on justifier la validité de cette conception fréquentiste en invoquant l'expression de la loi des grands nombres ?

Ce serait tomber dans une même ambiguïté qui mêle probabilité mathématique et "chances" estimées par un sujet observateur d'une expérience aléatoire. L'hypothèse énoncée par Hélène Ventsel est un fait d'observation d'un phénomène naturel que l'on pourrait qualifier de "loi du hasard" comme on qualifie de lois certains phénomènes physiques (loi de la pesanteur...).

Par contre la "loi des grands nombres" est un théorème de la théorie des probabilités développée à partir du modèle abstrait. Il est remarquable que ce théorème est bien illustré (jusqu'ici) par les faits d'observation, ce qui réconforte le probabiliste quant à la validité pratique de son modèle.

Nous voici donc placés devant deux approches de la probabilité, qui semblent contradictoires et complémentaires :

— une valeur *a priori* qui exploite les informations détenues par le sujet, relatives aux conditions de l'expérience aléatoire et à ses symétries, autorisant une hypothèse d'équiprobabilité quelque part,

— et une évaluation *a posteriori*, à partir des fréquences observées, de caractère objectif car attachée uniquement à l'événement.

L'enseignement des probabilités au Lycée doit faire la synthèse de ces deux approches conceptuelles.

Pour éclaircir le débat épistémologique, il n'est pas inutile de questionner l'Histoire, car cette dialectique l'a traversée ; elle a accompagné le développement du calcul des probabilités de sa naissance jusqu'à notre époque.

Je vous invite à cette petite promenade, soulignant quelques moments clés de l'élaboration du savoir probabiliste.

II — Histoire et épistémologie

1. Apparition d'une démarche scientifique

Dans *La République* (Livre... VII), sous la forme d'un dialogue entre Socrate et Glaucon (rapporté par Jean Dhombres [21, pp. 22 et 23]), Platon passe en revue les mathématiques :

L'Arithmétique, "*donne à l'âme un puissant élan vers la région supérieure, et la force à raisonner sur les nombres eux-mêmes, sans jamais souffrir qu'on introduise dans ses raisonnements des nombres qui représentent des objets visibles ou palpables* ". L'Arithmétique est la science par excellence et appartient toute entière au domaine des Idées.

La Géométrie, "*si elle nous oblige à contempler l'essence, elle nous convient ; si elle se borne à ce qui naît, elle ne nous convient pas*". La Géométrie a donc pour fonction de construire des modèles abstraits, permanents, qui rendent compte des propriétés des objets du monde sensible.

L'Astronomie, science de l'observation du ciel, étudie cette partie du monde sensible organisée en cycles stables, et fait appel pour cela à la géométrie des solides.

La Musique, contre partie de la Gymnastique, accorde l'harmonie des sens, mais n'y voyant "*que les œuvres mécaniques* ". Socrate la déclare impropre à "*attirer l'âme de ce qui naît à ce qui est* ".

Dans son inscription célèbre, Platon désigne sous le terme de géomètre, celui qui, tout en regardant le monde sensible, élève son raisonnement dans le monde des Idées pour atteindre la connaissance. Il englobe toutes les activités mathématiques qui se déroulent dans le monde des Idées.

Aristote (384-322 av. J.-C.), l'élève fameux de Platon, attribuera à la connaissance scientifique le rôle d'explication du monde réel. Sa pensée matérialiste le conduit à la recherche d'une description aussi fine que possible de la nature et de sa compréhension.

Aristote affirme : "*savoir, c'est connaître par le moyen de la démonstration*" (4). Il distingue les objets de la nature avec leurs propriétés à découvrir et les modèles théoriques à inventer, avec leurs termes primitifs, leurs définitions, leurs axiomes, leurs hypothèses et leurs théorèmes établis au moyen de démonstrations [2].

Mais les modèles qui étudient des concepts abstraits ne garantissent pas l'existence des objets matérialisant ceux-ci. Il revient au physicien d'en contrôler leur pertinence en les confrontant aux propriétés et lois naturelles qu'il a pu observer : "*pour parvenir à ces définitions, prémisses indispensables d'un syllogisme, le raisonnement inductif partant du sensible est le seul outil possible*" [21, p. 71]

C'est dans cet esprit que Pascal, 20 siècles après, désigne le calcul des probabilités par les termes de "*géométrie du hasard*", on dirait donc aujourd'hui "*mathématique du hasard*". Cette dernière étudie les phénomènes aléatoires au moyen d'outils mathématiques opérant dans le domaine des idées. Pour cela, il faut créer une théorie abstraite, avec ses concepts, ses définitions et ses théorèmes, sensés nous éclairer et nous permettre d'anticiper sur l'observation.

Tels les géomètres avec leurs "*figures qu'ils* ***modèlent*** *ou dessinent, [...], qu'ils emploient comme si c'étaient aussi des images, pour arriver à voir ces objets supérieurs qu'on n'aperçoit que par la pensée*" (*La République*, Livre VI, cité par Dhombres [21, p. 26]), les probabilistes exerceront leur art avec leurs expériences aléatoires, mais raisonneront dans leur modèle avec leurs concepts, leurs définitions et leurs théorèmes.

2. Naissance du calcul des probabilités

a) L'origine.

Aristote a tenté de cerner l'idée de hasard comme phénomène objectif [23, p. 211]. Il donne comme exemple la rencontre inopinée sur l'Agora entre un créancier et son débiteur. Pour lui le hasard se manifeste par la rencontre de deux processus indépendants, *à condition que l'on ait intérêt à remarquer cette rencontre.*

Les commerçants, banquiers et juristes des 14e et 15e siècles devaient intégrer les éventualités issues du hasard de la vie dans leurs "*calculs empiriques d'assurance d'une cargaison ou lors du partage des intérêts en fonction de l'argent placé dans des associations*" [16, p. 211].

On fait cependant remonter à la correspondance (1654) entre Blaise Pascal et Pierre Fermat [4, p. 43-49] l'idée moderne de probabilité en tant que mesure du degré d'incertitude d'un événement aléatoire, bien que Jérôme Cardan (1547) ait utilisé de telles proportions pour étudier les jeux de hasard [3].

C'est le problème célèbre des partis (5) (comment répartir les mises — faire le parti — entre

(4) Cité par A. Dahan-Dalmedico et J. Peiffer dans *Une histoire des mathématiques, routes et dédales* ed. Seuil. Points Sciences p. 54

(5) On pourra consulter sur ce sujet l'étude détaillée publiée dans la brochure de l'IREM de Strasbourg (Octobre 1992) : *Enseigner les probabilités en classe de première (programmes 1991)*

joueurs dans un jeu de hasard à plusieurs parties interrompu avant sa fin) qui a conduit Pascal et Fermat à la fameuse formule de la probabilité d'un événement A :

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Comprenant la généralité de cette définition, Pascal en dégage la nature à la lumière des idées de Platon et Aristote :

"Ainsi, joignant la rigueur des démonstrations de la science à l'incertitude du sort, et conciliant ces deux choses en apparence contradictoires, elle peut, tirant son nom des deux, s'arroger à bon droit ce titre stupéfiant : «La géométrie du hasard»." (6)

b) L'œuvre de Jacques Bernoulli : Ars Conjectandi *(1713, [5]).*

La conception pascalienne suppose l'équiprobabilité des événements élémentaires. Bernoulli en fait la remarque [5, p. 40] :

"on en est ainsi venu à ce point que pour former selon les règles des conjectures sur n'importe quelle chose, il est seulement requis d'une part que les nombres de cas soient soigneusement déterminés, et d'autre part que soit défini combien les uns peuvent arriver plus facilement que les autres ".

Cela s'applique à des situations aléatoires simples, comme les jeux de hasard qui sont conçus sous cette hypothèse précisément pour permettre le calcul des "chances" respectives de chaque joueur, *"de manière à se ménager l'équité "* selon les mots de Bernoulli [5, p. 40].

Mais c'est là une restriction majeure qui exclut toute application réelle aux phénomènes naturels complexes comme : l'apparition d'une maladie ou les phénomènes météorologiques ou encore la prévision des stratégies choisies par des joueurs dont les comportements sont imprévisibles.

En citant ces exemples [5, p. 42], Bernoulli met en cause l'idée qu'une probabilité soit calculée a priori, à partir des données de l'expérience aléatoire, exploitant les symétries du dispositif expérimental.

Il souligne qu'un dé a une chance sur six de tomber sur une face désignée à l'avance, caractère certes objectif de cette probabilité. Cependant, dans cette détermination *a priori*, il conteste le rôle possible de l'observateur dans la conduite de son calcul à partir de son interprétation des conditions de l'expérience et de ses propres informations : et si le dé était pipé ?

Il rejette alors cette conception *subjectiviste* de la probabilité pour proposer de la déterminer *a posteriori*, après observation d'un grand nombre d'expériences semblables [5, p. 42] :

"Mais à la vérité ici s'offre à nous un ***autre chemin*** *pour obtenir ce que nous cherchons. Ce qu'il n'est pas donné d'obtenir a priori l'est du moins a posteriori, c'est à dire qu'il sera possible de l'extraire en observant l'issue de nombreux exemples semblables; car on doit présumer que, par la suite, chaque fait peut arriver et ne pas arriver dans le même nombre de cas qu'il avait été constaté auparavant, dans un état de choses semblables, qu'il arrivait ou n'arrivait pas."*

(6) Cité en exergue par Jean-Louis Boursin in *Les structures du hasard, les probabilités et leurs usages*, ed Seuil, Points Sciences p. 5.

Bernoulli assimile probabilité d'un événement attendu et fréquence observée de cet événement. Cette probabilité est alors une donnée *objective* attachée à l'événement. Il justifie cette assimilation par la loi des grands nombres qu'il démontre ensuite en 12 pages à partir de la seule formule du binôme de Newton.

Ces pages véritablement géniales, fondent ce que nous avons appelé l'*approche fréquentiste* des probabilités, celle que les programmes du second degré d'aujourd'hui mettent en avant.

La difficulté réside alors dans l'*estimation* (et non le calcul) des probabilités élémentaires, elles-mêmes entachées d'incertitudes. C'est là "*un des points les plus délicats de la théorie des hasards*" selon Laplace [9, p. 38], cette estimation est un des objets principaux de la statistique inférentielle developpée au vingtième siècle.

Notons au passage que dans le même ouvrage, dès le premier chapitre, Bernoulli exprime à propos du hasard une pensée clairement déterministe, même si, époque oblige, elle doit s'accommoder d'une référence divine [5, p. 14] :

"*Tout ce qui bénéficie sous le soleil de l'être ou du devenir, passé, présent ou futur, possède toujours en soi et objectivement une certitude totale. C'est évident du présent et du passé : ce qui est ou a été ne peut pas ne pas être ou avoir été. Sur le futur il n'y a pas à discuter...*",

Bernoulli feint d'ignorer que c'est pourtant un problème majeur, pour faire sa révérence théologique :

"*...cependant ce n'est pas par la nécessité de quelque destin qu'il ne peut pas ne pas advenir, mais en raison soit de la prescience soit de la prédétermination divine*", et il énonce ce principe déterministe [5, p. 16] :

"car si n'arrivait pas avec certitude tout ce qui est futur, on ne voit pas comment le Créateur suprême pourrait conserver entière la gloire de son omniscience et de son omnipotence ".

Enfin Bernoulli tente d'échapper au débat philosophique sur le libre arbitre de l'Homme, si bien développé un siècle plus tard par Denis Diderot dans *Jacques le fataliste* (1796), en ajoutant :

"*Quant à dire comment cette certitude de l'avenir peut subsister avec la contingence ou la liberté des causes secondes, que d'autres en disputent; pour nous, nous ne voulons pas toucher aux points étrangers au but que nous visons*".

Dans cette conception déterministe la probabilité apparaît comme ***"un degré de la certitude et en diffère comme la partie diffère du tout ".*** [5, p. 16].

c) Bayes ou le glissement subjectiviste
La doctrine des chances *(1763, [6])*

Thomas Bayes s'est très tôt (1731) intéressé à l'estimation des probabilités des événements naturels, suivant en cela les orientations de Bernoulli. Cependant, il décrète l'équiprobabilité a priori des valeurs possibles pour une probabilité inconnue, quitte à la réajuster a posteriori, introduisant ainsi deux notions de probabilité : la première, que l'on cherche à estimer, est objective, la seconde, appréciant la valeur possibles de la précédente, posée a priori, est subjective.

Dans sa *petite histoire du calcul des probabilités*, Bernard Bru précise [17, p. 147] :

" [Selon Bayes], *la probabilité d'un évènement est semblable à une mesure physique sur laquelle on ne disposerait a priori d'aucune information* [...]. *Il est donc naturel de supposer que toutes les valeurs de cette probabilité sont, a priori, «également probables»* [...]. *Si ensuite, on observe les réalisations de l'événement au cours de n épreuves, la probabilité de cet événement va changer, puisqu'on dispose maintenant d'informations objectives*".

Et Bernard Bru fait le commentaire qui suit :

"Bayes, pour la première fois, va laisser déraper la probabilité-objet mathématique que la géométrie du hasard a extrait de la description probabiliste vers la probabilité au sens du dictionnaire Larousse : à une notion subjective (toutes les informations qui fondent notre intime conviction) il associe un objet mathématique (formel) : une loi de probabilité, et en utilise les propriétés pour remonter au réel.

Formellement, mathématiquement, la méthode de Bayes permet de résoudre le problème de l'estimation des probabilités ; mais quelle est la valeur pratique du résultat obtenu, son adéquation à l'univers (extérieur à notre subjectivité) ?"

De fait, Bayes lui-même n'était pas très sûr de ses résultats et ne les publia pas (sa "doctrine des chances" sera publiée deux ans après sa mort).

Bernard Bru ajoute :

"Ce glissement de l'objectif au subjectif à travers le miroir de la géométrie du hasard, bien qu'aperçu par Condorcet, va s'opérer en douceur et Laplace maniera les deux types de probabilités sans crise de conscience particulière. Ce sera le mérite de Cournot (7) *"d'attirer l'attention sur le problème, 50 ans plus tard"*.

Cependant, les méthodes bayésiennes, issues de cette idée d'attribuer une probabilité aux causes d'un événement qui s'est réalisé à partir d'une valeur prise a priori et recalculée en fonction de l'observation, ont leur domaine de validité et d'efficacité. Elles recueillent aujourd'hui des adeptes enthousiastes et des détracteurs systématiques.

Notons que la démarche qui considère que l'événement observé était a priori le plus probable, est à la source de la méthode du maximum de vraisemblance utilisée pour déterminer des estimateurs en statistique inférentielle.

Mais Bayes introduit bel et bien une nouvelle sorte de probabilité (appelée subjective), mathématiquement identique à la probabilité de la géométrie du hasard, elle correspond cependant à la géométrie d'un jeu de hasard qui ne serait joué qu'une seule fois. Elle ne peut convenir à l'approche fréquentiste de ce concept.

d) Les hésitations de D'Alembert : article Croix ou pile *de l'Encyclopédie [7]*

La loi des grands nombres (sous la forme du théorème de Bernoulli) exprime qu'au cours d'un grand nombre d'expériences aléatoires répétées dans des conditions identiques, la fréquence d'un événement tend vers sa probabi-

(7) Antoine Augustin Cournot (1801-1877), économiste, philosophe et mathématicien, chercha, comme Auguste Comte, à établir une classification des connaissances humaines. Son œuvre mathématique concerne principalement les probabilités [10]. Reprenant l'étude de l'idée de hasard du point de vue objectif d'Aristote, il écrit *"les événements amenés par la combinaison ou la rencontre d'autres événements qui appartiennent à des séries indépendantes les unes des autres sont ce qu'on nomme des résultats du hasard"*

lité (plus exactement, il n'y a aucune chance d'observer le contraire)

Cette propriété du hasard n'est pas aussi intuitive qu'il y paraît. Cela suppose en particulier que les n premières issues observées n'ont aucune influence sur la réalisation de la n+1ème. Or nombreuses sont les personnes qui, avant de jouer au loto, consultent les résultats des tirages précédents.

On rencontre à cet égard deux types de raisonnements : si une boule est sortie plus souvent que les autres, par une fausse interprétation d'une sorte de loi des séries, certaines préfèrent parier sur cette boule ; d'autres par contre, sur la base d'une conception erronée de la loi des grands nombres, estiment que la nature devrait rétablir l'équilibre lors des tirages suivants et ne parient pas sur cette même boule.

Jean Le Rond D'Alembert s'est interrogé en son temps (1760) sur l'attitude à adopter dans une série de piles ou faces (croix ou piles à son époque) qui montrerait une forte proportion de "piles".

D'Alembert réfute l'indépendance de deux coups consécutifs, il estime au contraire que l'occurrence d'une répétition rend plus probable la réalisation d'un événement différent (8) :

"dans le cours ordinaire de la nature, le même événement (quel qu'il soit) arrive assez rarement deux fois de suite, plus rarement trois et quatre fois, et jamais cent fois consécutives " .

(8) Correspondance de d'Alembert (10e mémoire, 1767), citée par Michel Paty dans son article "La critique par d'Alembert des conditions d'une théorie des probabilités physiques", *in Fundamenta Scientiae*, 1987

Mais il fait une différence entre les répétitions en petit nombre, qui peuvent être l'effet du hasard, et celles de grande fréquence, qui, si on les observe, doivent être rapportées à une cause : dans une longue série de piles, une attitude raisonnable ne serait-elle pas au contraire d'inférer que la pièce est truquée et que les observations suivantes vont confirmer le déséquilibre entre les piles et les faces. Il est alors préférable de proposer une valeur (d'origine fréquentiste) de la probabilité des piles supérieure à 0,5 (9):

"Ce qu'on vient de dire est fondé sur la supposition que pile ne soit pas arrivé de suite un très grand nombre de fois : car il serait plus probable que c'est l'effet de quelque cause particulière dans la construction de la pièce, et pour lors, il y aurait de l'avantage à parier que pile arriverait encore".

Déjà Pascal (cité par D'Alembert) avait soulevé le paradoxe : *"Le cas dont il s'agit* [avec trois dés, faire 20 fois de suite trois six] *est **physiquement** possible ? Il faut dire qu'il ne l'est pas; quoiqu'il soit possible **mathématiquement**".*

Cette non indépendance entre des tirages successifs s'exprime également dans l'attribution de la valeur 2/3 à la probabilité d'obtenir un "face" en deux lancers d'une pièce. D'Alembert explique dans l'*Encyclopédie* (article "Croix ou pile") :

"il n'y a proprement que trois combinaisons de possibles : croix au premier coup; pile, croix aux premier et second coup; pile, pile aux deux coups. Donc il y a 2 contre 1 à parier ".

Cette hypothèse d'équiprobabilité globale

(9) Ibid

appliquée à l'ensemble des résultats observables (car si on a amené croix au premier coup, la partie s'arrête) est source de nombreux comportements erronés en probabilités. D'Alembert l'écrit explicitement dans l'article *"gageure"* de l'*Encyclopédie* :

"...il me paraît toujours difficile de bien expliquer pourquoi et comment l'avantage peut être triple, lorsqu'il n'y a que deux coups favorables".

En réalité, D'Alembert refuse de se placer dans le domaine des Idées de Platon, il limite le calcul des probabilités à un statut de description de la réalité sensible, tout en y recherchant des preuves mathématiques. Il refuse d'envisager abstraitement une longue série de piles sans que la pièce soit truquée, il ne veut pas imaginer la poursuite fictive d'une partie de piles ou faces interrompue par l'arrivée de l'événement attendu. Cette confusion entre réalité sensible et modèle mathématique a empêché ce mathématicien philosophe d'y voir clair en probabilités.

3. Le développement mathématique du calcul des probabilités

a) Laplace fait le ménage : L'essai philosophique sur les probabilités *(1825)*

Il revenait à Pierre-Simon Laplace de clarifier le calcul des probabilités en le plaçant résolument dans le cadre mathématique. Dans son introduction à l'essai philosophique, reprenant le déterminisme de Bernoulli en y remplaçant le "Créateur" par la "nature" (10), il règle le débat sur la réalité sensible du hasard par l'affirmation d'un déterminisme absolu.

Après avoir invoqué *"le principe de la raison suffisante"* selon lequel *"une chose ne peut pas commencer d'être, sans une cause qui la produise"* [9, p. 32], il a cette formule célèbre, véritablement géniale par sa concision :

"Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre",

puis il ajoute [9, p. 33] :

"Une intelligence qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome ; rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux".

Dans ces conditions [9, p. 34], *"la probabilité est relative en partie à notre ignorance, en partie à nos connaissances"*, affirmant ainsi une conception subjectiviste, confirmée par la reprise de la définition de Pascal de la probabilité comme donnée de base d'un modèle mathématique limité. Il énonce [9, p. 35] :

"La théorie des hasards consiste à réduire tous les événements du même genre à un certain nombre de cas également possibles, c'est à dire tels que nous soyons également indécis sur leur existence, et à déterminer le nombre de cas favorables à l'événement dont on cherche la probabilité".

Laplace construit alors son modèle mathé-

(10) Le matérialisme de Laplace s'est exprimé dans cette réponse à l'empereur qui l'interrogeait sur l'existence de Dieu : *"Sire je n'ai pas eu besoin de cette hypothèse"* (cité par René Thom [9, p. 27])

matique sur la base de dix principes, rassemblant définitions et axiomes traduisant son approche pascalienne. Les deux premiers sont déterminants [9, p. 38] :

"Premier principe .

[La probabilité] *est le rapport du nombre des cas favorables à celui de tous les cas possibles.*

Deuxième principe.

Mais cela suppose les divers cas également possibles."

(Laplace exprime ici clairement les limites de son modèle, mais il laisse la porte ouverte à des problèmes plus généraux). Il ajoute :

"S'ils ne le sont pas, on déterminera d'abord leurs possibilités respectives dont la juste appréciation est un des points les plus délicats de la théorie des hasards". Et il énonce l'axiome : *"Alors la probabilité sera la somme des possibilités de chaque cas favorable ".*

En application de ces deux principes, Laplace reprend l'exercice de D'Alembert "croix" ou "pile" en deux lancers [9, p. 39] :

"On peut ne compter à ce jeu que trois cas différents. [...] . Cela réduirait la probabilité (d'avoir croix au moins une fois en deux coups) à 2/3, si l'on considérait avec D'Alembert ces trois cas comme également possibles ".

Laplace démontre alors qu'ils ne le sont pas et conclut :

"On a 3/4 pour la probabilité cherchée, ce qui s'accorde avec ce que l'on trouve dans la supposition où l'on joue les deux coups. Cette supposition ne change point le sort de celui qui parie pour cet événement : elle sert seulement à réduire les divers cas à des cas également possibles ".

Mais obnubilé par cette équiprobabilité nécessaire, Laplace en illustration de son sixième principe définissant la probabilité conditionnelle, en vient à en faire indûment l'hypothèse implicite dans un problème où elle n'a aucune raison d'être, deux pages après avoir épinglé D'Alembert [9, pp. 42 à 45] :

*"**Sixième principe** [...] la probabilité de l'existence d'une quelconque des causes auxquelles un événement observé peut être attribué, est donc une fraction dont le numérateur est la probabilité de l'événement, résultante de cette cause, et dont le dénominateur est la somme des probabilités semblables relatives à toutes les causes ."*

Laplace suppose ici égales les probabilités a priori. Avec deux cause C_1 et C_2 de l'événement E, on peut traduire son texte par la formule suivante, donnée ici avec les notations contemporaines :

$$P(C_i/E) = \frac{P(E\,C_i)}{P(E\,C_1) + P(E\,C_2)}$$

mais il ajoute :

"si diverses causes considérées a priori *sont inégalement probables, il faut au lieu de la probabilité de l'événement, résultante de chaque cause, employer le produit de cette probabilité, par la possibilité de la cause elle-même ".*

Il donne ainsi la formule improprement

dite de Bayes :

$$P(C_1/E) = \frac{P(E/C_1)P(C_1)}{P(E/C_1)P(C_1)+P(E/C_2)P(C_2)}$$

En guise d'application, Laplace propose l'exercice suivant:

"Imaginons une urne qui ne renferme que deux boules dont chacune soit ou blanche, ou noire. On extrait une de ces boules, que l'on remet ensuite dans l'urne, pour procéder à un nouveau tirage. Supposons que dans les deux premiers tirages on ait amené des boules blanches; on demande la probabilité d'amener encore une boule blanche au troisième tirage".

Et Laplace applique sa première formule pour donner le résultat : 9/10, sans justifier l'hypothèse de l'égalité des probabilités a priori des deux urnes possibles [blanche, blanche] et [blanche, noire]. Sa solution ne correspond à aucune hypothèse de son énoncé et élimine les cas où l'une des deux urnes aurait été tirée avec un dé par exemple.

Y aurait-il sous-jacente cette conception subjectiviste, déjà présente chez Bayes, qu'en l'absence de toute indication sur les différentes possibilités des résultats d'une expérience aléatoire, l'attitude la plus raisonnable est de proposer l'équiprobabilité ?

Ou y aurait-il (ce qui ne serait pas acceptable aujourd'hui de la part d'un élève de première !) l'introduction d'une probabilité 1/2 pour chaque urne en l'absence d'expérience aléatoire ?

L'approche fréquentiste suppose la réalisation matérielle de cette situation et met en évidence la nécessité du choix aléatoire entre les deux urnes par tel ou tel dispositif, elle souligne alors l'imprécision de l'énoncé de Laplace et apporte une réponse claire à ce problème.

b) Poincaré et le déterminisme

Le 19e siècle voit se développer le calcul des probabilités, fondé par Laplace dans sa "théorie analytique des probabilités" [9'] publiée en 1812. De nombreux résultats nouveaux sont obtenus par Gauss, Poisson, Bienaymé... L'analyse a fait aussi de grands progrès avec Fourier, Riemann... Mais le calcul des probabilités se heurtait aux mêmes paradoxes que le reste des mathématiques utilisant largement l'infini et les passages à la limite insuffisamment fondés, comme le souligne le paradoxe de Bertrand proposé en 1899 (*cf.* la note 2 et l'article "Paradoxes et lois de probabilités" paru dans *Repères-IREM* n° 13).

De plus, la complexité des phénomènes naturels que l'on cherchait à modéliser de plus en plus finement, finit par faire douter du déterminisme laplacien, trop simple pour y être adapté. Évoquant la théorie cinétique des gaz, Henri Poincaré déclare, dans l'introduction de son cours de physique mathématique à la faculté des sciences de Paris [12, pp. 3 et 4] :

"Il faut donc bien que le hasard soit autre chose que le nom que nous donnons à notre ignorance [...]. *Et pour les phénomènes fortuits eux-mêmes, il est clair que les renseignements que nous fournit le calcul des probabilités ne cesseront pas d'être vrais le jour où ces phénomènes seront mieux connus".*

Et Poincaré donne cette définition du hasard, après avoir proposé l'exemple du cône posé sur sa pointe de telle sorte qu'on ne peut prévoir de quel côté il va tomber :

"Une cause très petite, qui nous échappe, détermine un effet considérable que nous ne pouvons pas ne pas voir, et alors nous disons que cet effet est dû au hasard".

Cette définition s'accommode du déterminisme, en réduisant le hasard à la manifestation d'un phénomène dissipatif ; Poincaré ajoute :

*"Si nous connaissions exactement les lois de la nature et la situation de l'univers à l'instant initial, nous pourrions prédire exactement la situation de ce même univers à un instant ultérieur. Mais, lors même que les lois naturelles n'auraient plus de secret pour nous, nous ne pourrions connaître la situation qu'**approximativement**. Si cela nous permet de prévoir la situation ultérieure avec la **même approximation**, c'est tout ce qu'il nous faut, nous disons que le phénomène a été prévu, qu'il est régi par des lois; mais il n'en est pas toujours ainsi, il peut arriver que de petites différences dans les conditions initiales en engendrent de très grandes dans les phénomènes finaux; une petite erreur sur les premières produirait une erreur énorme sur les derniers. La prédiction devient impossible et nous avons le phénomène fortuit".*

Poincaré reprend ici une formulation du déterminisme analogue à celle de Laplace, mais il adopte une attitude objectiviste [12, p. 16] :

"Le hasard étant ainsi défini dans la mesure où il peut l'être, a-t-il un caractère objectif ? On peut se le demander. J'ai parlé de causes très petites ou très complexes"

et Poincaré répond par l'affirmative [12, p. 18] : "[Le caractère de complexité] *est donc relatif, s'il conserve son caractère objectif, c'est parceque tous les hommes ont à peu près les mêmes sens, que la puissance de leurs instruments est limitée"*.

Dans *la science et l'hypothèse* [12', p. 195], il clarifie cette opposition entre positions subjectivistes et objectivistes :

*"un joueur veut tenter un coup; il me demande conseil. Si je le lui donne, je m'inspirerai du calcul des probabilités, mais je ne lui garantirai pas le succès. C'est là ce que j'appellerai **la probabilité subjective.** [...]*
*Mais je suppose qu'un observateur assiste au jeu, qu'il en note tous les coups et que le jeu se prolonge longtemps; quand il fera le relevé de son carnet, il constatera que les événements se sont répartis conformément aux lois du calcul des probabilités. C'est ce que j'appellerai **la probabilité objective,** et c'est ce phénomène qu'il faudrait expliquer".*

Poincaré cite encore l'exemple de la météorologie. Ses travaux en mécanique céleste mettent en évidence l'instabilité de systèmes complexes dont l'évolution dépend de manière très sensible des conditions initiales. Il est ainsi à l'origine des conceptions modernes sur le hasard et le statut des probabilités.

c) Puis vint Kolmogorov : les fondements mathématiques du calcul des probabilités

Après Poincaré, il restait à exploiter les avancées les plus récentes de l'analyse mathématique, Georges Cantor et David Hilbert en ont clarifié les fondements avec la théorie des ensembles, particulièrement adaptée à la description des modèles probabilistes. Avec les travaux d'Emile Borel (1898, [14]) en théorie de la mesure et ceux de Henri Lebesgue (1904, [13]) en intégration, l'outil mathématique du calcul des probabilités devenait disponible, dès lors que l'on *reconnaissait à la probabilité les propriétés d'une mesure.*

Paul Lévy et Khintchine avaient développé parallèlement la théorie des variables aléatoires, Andréi Kolmogorov [15] opérera cette identification et jettera les fondements de la théorie moderne des probabilités, lui donnant toute sa dimension mathématique. C'est cette théorie, dans sa présentation élémentaire, qui fait l'objet des cours actuels dans les universités.

Dans ce cadre, quelles que soient nos conceptions philosophiques et épistémologiques sur le hasard, objectivistes ou subjectivistes sur les probabilités, quelle que soit notre approche du concept de probabilité, pascalienne ou fréquentiste, elles sont compatibles avec l'exposé moderne du calcul des probabilités, qui évite soigneusement ces débats en développant une théorie axiomatique susceptible de recevoir de nombreuses interprétations.

La théorie de Kolmogorov a eu l'avantage de rendre plus clair le statut de la définition de Pascal : *elle revient au choix de la mesure uniforme sur un ensemble fini d'événements élémentaires.*

d) Avancées contemporaines : déterminisme et complexité, probabilités et statistiques

L'enseignement des probabilités est resté longtemps basé sur cette approche pascalienne, reprenant les exposés de Laplace, de Poincaré, exploitant largement les modèles d'urnes qui rendent compte des situations où l'ensemble des événements élémentaires est fini (avec équiprobabilité), explorant les lois principales dégagées au 19e siècle et développant l'outil mathématique.

Aujourd'hui cet enseignement se heurte à ses limites déjà soulignées par Bernoulli. Dès lors qu'il s'agit de développer les applications dans tous les domaines d'activité des sociétés humaines, de l'économie à la technologie, de la statistique à la biologie, les conceptions subjectivistes sont inopérantes.

Il convient de préparer les élèves à l'utilisation du calcul des probabilités dans des situations concrètes les plus diverses, notamment dans le cadre de l'inférence statistique, estimations et tests d'hypothèses, développés au début du 20e siècle.

L'approche fréquentiste est devenue nécessaire, le passage par la modélisation est essentiel.

Ce point de vue est central dans la thèse de didactique des probabilités de Jacques Bordier (1991, [34]) qui cite en introduction cette remarque de B. Gnedenko et A. Kolmogorov :

"La valeur épistémologique de la théorie des probabilités est basée sur ceci : dans leur action collective les phénomènes aléatoires, à large échelle, créent des régularités strictes et non aléatoires. Le concept même de la probabilité mathématique serait sans utilité, s'il ne trouvait sa concrétisation dans la fréquence d'arrivée d'événements, suite à des expériences nombreuses réalisées dans des conditions uniformes."

La pensée déterministe reste cependant prégnante dans toute démarche scientifique : *"Dieu ne joue pas aux dés"* selon la formule d'Einstein, reprise par René Thom [9, p. 23] qui ajoute :

"dans ce conflit déterminisme-hasard, la Science est déterministe pour des raisons de principe".

Mais aujourd'hui, on s'interroge sur l'existence du hasard dans les phénomènes naturels eux-mêmes, quel que soit le degré de précision possible dans leur observation. Induites par la relation d'incertitude de Heisenberg en mécanique quantique, ces questions sont développées à notre époque par Jacques Monod dans *Le hasard et la nécessité* (11), puis par Ilya Prigogine et Isabelle Stengers dans *La nouvelle alliance* [24] et par les théoriciens de la complexité comme Edgar Morin sur le plan philosophique et épistémologique (12), David Ruelle pour les théories du chaos [25] et Ivar Ekeland dans son merveilleux livre *Au hasard* [22] (pour sa présentation, on pourra relire la note de lecture de Jean-Claude Duperret parue dans *Repères-IREM* n° 10).

Les écrits contemporains sur le chaos, le déterminisme, le hasard et la complexité sont maintenant très nombreux. Citons particulièrement : *Le hasard aujourd'hui* [23], recueil d'émissions de France Culture animées par Emile Noël, questionnant sur le hasard un ensemble riche de spécialistes de disciplines variées.

III — Modélisation mathématique des phénomènes aléatoires

1. L'approche fréquentiste dans l'enseignement universitaire

L'ouvrage de base pour cet enseignement universitaire est le livre *Calcul des probabilites* d'Alfred Rényi, publié en 1966 chez Dunod [28], réédité par Jacques Gabay.

Dans l'introduction de ce cours, Rényi énonce [28, p. 25] :

"Nous appellerons ***probabilité*** *d'un événement le nombre autour duquel oscille la fréquence relative de l'événement considéré. La probabilité d'un événement peut changer avec les conditions dans lesquelles est faite l'épreuve. Le fait que la fréquence relative* (13) *presente vraiment une stabilité et la valeur numérique autour de laquelle oscillent les observations statistiques ne peuvent être déterminés que par l'expérience".*

Rényi enfonce le clou [28, p. 26] :

"Nous considérons donc la probabilité comme une valeur indépendante de l'observateur, qui indique approximativement avec quelle fréquence l'événement considéré se produira au cours d'une longue série d'épreuves".

Puis Rényi revient sur les deux conceptions historiques de la probabilité [28, p. 26] :

"Dans la vie quotidienne on exprime souvent un jugement subjectif sur les possibilités de réalisation d'un événement aléatoire. Mais la théorie mathématique des probabilités ne s'occupe pas de ces jugements subjectifs , elle concerne les probabilités objectives, qui peuvent être mesurées comme des grandeurs physiques [...] Mais il faut soigneusement distinguer les deux notions : de probabilité -qui est un nombre fixe -, et de fréquences relatives -qui dépendent du hasard".

Enfin Rényi, après avoir démontré la loi des grands nombres, justifie la cohérence de l'approche fréquentiste [28, p. 144] :

(11) Jacques MONOD : *Le hasard et la nécessité*, essai sur la philosophie naturelle de la biologie moderne- éd. Seuil, Paris 1970.

(12) Edgar MORIN : *Science avec conscience*, éd. Seuil, Points Sciences 1990.

(13) Le terme de fréquence (des réalisations d'un événement) remplace dans nos programmes scolaires ce que Rényi appelle ici fréquence relative.

"Lors de l'introduction du concept de probabilité, nous avons assigné une probabilité à des événements dont la fréquence relative, au cours d'une longue série d'épreuves, manifestait une certaine stabilité. Ce fait, la stabilité de la fréquence relative, vient d'être démontré mathématiquement. Il est remarquable que la théorie rende possible une description précise de cette stabilité ; cela témoigne sans aucun doute en faveur de sa puissance.

Il semblerait qu'il s'agisse alors d'un cercle vicieux. Nous avons en effet défini la probabilité grâce à la stabilité de la fréquence relative, mais d'autre part la notion de probabilité intervient pour caractériser cette stabilité. En réalité, il s'agit pourtant de deux choses entièrement différentes. La « définition » de la probabilité comme valeur autour de laquelle oscille la fréquence relative n'est pas une définition mathématique mais une description du substrat concret du concept de probabilité. La loi des grands nombres de Bernoulli par contre est fondée sur la définition mathématique de la probabilité et par conséquent il n'y a là aucun cercle vicieux".

Ainsi Rényi fait clairement la distinction entre la description du monde sensible et le modèle mathématique abstrait, servant de cadre aux démonstrations.

2 . Introduction du modèle probabiliste en première

L'enseignement des probabilités n'a pas pour objectif de remplacer des conceptions par d'autres ou de réfuter telle ou telle approche. Comme le souligne l'ex-futur programme de la série L, il propose de construire un outil mathématique par la modélisation d'expériences aléatoires dont la description en est une étape essentielle.

La probabilité devient alors un concept abstrait, mathématique, selon les termes de Kolmogorov, ce qui la dégage des dialectiques subjectif-objectif, a priori-a posteriori, ou géométrique-fréquentiste, tout en permettant le contrôle de son adéquation aux situations concrètes par l'ajustement des hypothèses de travail et des données numériques de base.

Cette démarche est bien mise en évidence dans la thèse de Jacques Bordier [34] qui propose d'exploiter à fond l'ordinateur pour faire travailler les élèves sur des simulations de situations aléatoires qui devront être ensuite modélisées

a) Une analogie

Dans son introduction, Rényi [28, p. 26] compare la probabilité à une mesure physique. Développons cette analogie avec la mesure des longueurs :

Problème du monde sensible

Expérience aléatoire.	Mesure de longueurs.
Présentation de l'expérience aléatoire : étude de la durée de vie d'une ampoule électrique.	Présentation de la situation : détermination de la hauteur d'une muraille.
Description de l'expérience : mise en évidence d'un phénomène rare et sans mémoire, satisfaisant aux conditions d'une situation de Poisson.	Description de la situation : mise en évidence des propriétés de parallélisme, satisfaisant aux conditions du théorème de Thalès.

Passage au modèle

Définition de l'ensemble Ω des durées possibles et introduction d'une variable aléatoire de type exponentiel. Etude des lois diverses intervenant dans ce modèle : loi de Poisson du nombre de pannes dans un intervalle de temps donné, lois gamma des durées de vie avant la $n^{ième}$ panne.	Représentation plane d'une configuration de Thalès en géométrie euclidienne et introduction d'une mesure des longueurs dans ce plan. Etude des propriétés de la configuration, détermination des rapports de longueurs pertinents et relations avec la longueur inconnue.

Retour au monde sensible : introduction d'un instrument de mesure

Comptage du nombre moyen de pannes dans un intervalle de temps donné au cours d'un grand nombre d'observations.	Mesure, au moyen d'une chaîne d'arpenteur, des dimensions accessibles sur le terrain.

Injection dans le modèle des données numériques

Détermination du paramètre de Poisson. Calcul des probabilités d'événements attachés à cette expérience, détermination de la durée moyenne de la vie d'une ampoule.	Calcul des rapports de Thalès. Calcul des longueurs intervenant dans la configuration et détermination de la longueur cachée.

Confrontation avec l'observation

Etude statistique de la durée de vie des ampoules électriques de même type, impact des paramètres non pris en compte dans le modèle.	Etude de la fiabilité de ce système de mesure dans différentes situations, impact des approximations dans les mesures.

b) Description d'une expérience aléatoire

Arrivés en classe de première, les élèves ont été confrontés depuis longtemps avec les phénomènes aléatoires, en particulier avec les jeux de hasard. Ils ont même pu émettre certaines inférences de nature probabiliste, sans bien pouvoir en mesurer la fiabilité.

matiques, on se propose d'étudier ces phénomènes et de les mesurer. Pour permettre aux élèves ce passage de la perception empirique à l'approche scientifique, on n'évitera pas un réel travail de fond sur la modélisation des expériences aléatoires.

A cette notion d'expérience aléatoire sont associées des hypothèses importantes :

Une expérience aléatoire doit conduire à un *semble de résultats possibles*, bien identifiés.

De plus, elle doit être *reproductible* dans les mêmes conditions (au moins par la pensée).

Ainsi, les faits historiques ne peuvent être considérés comme résultant d'expériences aléatoires et ne sauraient être probabilisés.

Mais les termes "mêmes conditions" conduisent à des difficultés : une pensée déterministe comme celle de Laplace contestera l'aléatoire des issues possibles si les conditions de l'expérience sont *rigoureusement* les mêmes.

Malgré cela, les résultats possibles ne peuvent être prévus à l'avance. Plus exactement, les conditions de l'expérience, *telles qu'elles peuvent être décrites et reproduites* ne déterminent pas l'un des résultats possibles de manière absolue.

Chacun des résultats possibles est alors le fruit du "hasard" se conjuguant avec les conditions de l'expérience, que ce hasard soit dû à notre ignorance ou qu'il soit réellement inscrit dans les phénomènes naturels ou humains étudiés.

La compréhension de cette notion d'expérience aléatoire n'est pas spontanée, d'autant que le terme d'"aléatoire" est utilisé sans discernement par les élèves.

Nous savons que le langage et les opérations sur les ensembles permettent effectivement une bonne représentation de ces notions d'expérience aléatoire, de résultat possible et d'événement.

Nous rencontrons là une difficulté didactique dans l'introduction du vocabulaire ensembliste : les programmes actuels des collèges et des lycées ne prévoient pas d'aborder ce vocabulaire et les schémas qui lui sont liés (patates et autres diagrammes) avant le chapitre des probabilités.

Ces outils élémentaires ne sont donc pas disponibles pour les élèves de première qui se heurtent à deux difficultés à la fois : comprendre le processus de modélisation et se familiariser avec les ensembles. Il serait donc judicieux que leur introduction soit plus explicite à l'occasion d'autres chapitres du cours de mathématiques.

Notons que le membre de phrase du programme précédent (TD et TC, 1986) :

"l'objectif est d'entraîner les élèves à organiser, grâce à un minimum de langage ensembliste, des données..."

disparaît de manière à ce que les opérations ensemblistes ne soient pas objet d'étude en elles-mêmes, ce qui répond à la critique de l'enseignement du formalisme ensembliste avec les glissements de sens auxquels il donnait lieu. Notons cependant que dans les connaissances exigibles, le nouveau programme indique :

"Les élèves doivent savoir calculer la probabilité de la réunion d'événements disjoints, d'un événement contraire,...".

c) Introduction des probabilités

Après avoir travaillé sur les combinaisons logiques d'événements que le modèle ensembliste permet de clarifier, les élèves peuvent poursuivre la construction du modèle en y introduisant les probabilités, comme le suggère l'analogie présentée ci-dessus.

Dans l'approche fréquentiste, la notion de modéle mathématique conduit à distinguer l'ensemble Ω, ensemble représentant les issues possibles de l'expérience aléatoire E, de ces issues elles-mêmes.

Elle conduit à faire la distinction entre la probabilité P(A) (qu'à la suite de Kolmogorov on pourrait qualifier de mathématique) d'un événement A, et la valeur éventuellement approchée p que l'on attribue à l'événement réel représenté par A.

P(A) est la mesure de la partie A de Ω, nombre interne au modèle, p est celui que l'on peut obtenir pour estimer cette probabilité, afin de rendre adéquat le modèle à la réalité décrite.

Cette valeur numérique peut alors être obtenue par deux moyens possibles :

— soit **a priori** si l'hypothèse d'équiprobabilité est acceptable dans le modèle, ce que des considérations expérimentales ou heuristiques permettent de justifier,

— soit **a posteriori** par l'étude de la stabilisation de la fréquence de A (il suffit d'ailleurs de prendre la fréquence obtenue à la suite du plus grand nombre possible d'expériences $\mathcal{E}$ répétées)

Ce choix est accompagné de la remarque essentielle : dans les cas d'équiprobabilité ces deux possibilités aboutiront à des résultats aussi voisins que l'on veut, la valeur "mathématique" retenue pour la probabilité étant celle que l'on peut *calculer* a priori, car elle traduit précisément l'hypothèse théorique d'équiprobabilité que l'on a choisi d'accepter.

Le lien entre probabilité et fréquence est alors clair [14], et les propriétés mathématiques des probabilités peuvent être aisément dégagées de celles des fréquences, comme cela est indiqué dans l'encadré suivant :

L'ensemble univers Ω décrivant l'expérience est constitué des r événements élémentaires ω_i représentant les issues possibles.

a- Les fréquences observées $f_{i,n}$ des w_i lors de la répétition n fois de l'expérience vérifient évidemment :

$$\sum_{i=1}^{r} f_{i,n} = 1 \text{ avec } 0 \leq f_{i,n} \leq 1, \text{ on doit donc avoir : } \sum_{i=1}^{r} p_i = 1, \text{ avec } 0 \leq p_i \leq 1.$$

b- La fréquence d'un événement A réalisé par plusieurs w_i est la somme des fréquences de chacun de ces w_i : $f_n(A) = \sum_{\omega_i \text{ réalisant } A} f_{i,n}$

En cohérence, dans le modèle probabiliste, on pose :

(14) Pour le développement de ce lien, on pourra également se reporter au bulletin de l'IREM de Besançon n° 45 (Juin 91) ou à l'article paru dans *Repères-IREM* n° 6 (Janvier 92).

$$P(A) = \sum_{\omega_i \in A} p_i \text{ , avec : } P(\{\omega_i\}) = p_i$$ [15].

c- On a les conséquences immédiates : $P(\Omega) = 1$; $P(\emptyset) = 0$;

$P(\overline{A}) = 1 - P(A)$; si $A \subset B$, alors $P(A) \leq P(B)$;
ainsi que la formule des probabilités totales : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Les relations inscrites dans l'encadré permettent de calculer les probabilités d'événements combinés, à partir de la donnée des probabilités élémentaires qui caractérisent le modèle, après diverses manipulations de logique ensembliste.

L'équiprobabilité apparaît alors comme un cas très particulier, la traduction mathématique de cette hypothèse se réduisant à l'égalité : $P(\{\omega_i\}) = 1/r$ pour tout i.

La formule de Pascal :

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

s'en déduit immédiatement et ouvre la porte aux problèmes de combinatoire dont l'étude plus détaillée est rejetée en terminale [16] où l'on poursuit l'initiation aux probabilités dans deux directions :

— L'introduction des outils de base que sont *les variables aléatoires* et les lois de probabilité,

— la notion de *probabilité conditionnelle.*

3 . Développements en terminale

a) Liaison première - terminale

Avant de donner du sens à l'étude des variables aléatoires et aux probabilités conditionnelles, il est bon de s'assurer que les élèves ont une bonne compréhension des notions d'expérience aléatoire et de probabilité.

Dans une terminale F de l'académie de Besançon, les réponses à un pré-test passé en janvier 1993 ressemblaient généralement à celle-ci :

"une expérience aléatoire est une expérience dont on ne sait pas trop d'avance si elle va marcher. Par exemple une interro de maths".

Dans la plupart des réponses, l'incertitude portait non pas sur les issues de l'expérience mais sur le déroulement de l'expérience elle-même. L'exemple d'expériences de physique qui tournent mal est également cité.

On peut s'interroger sur le traitement réservé au chapitre sur les probabilités dans certaines classes de premières en 91-92. Rejeté souvent en dernière semaine ou pas traité du

(15) Pour une présentation détaillée du modèle probabiliste issu de la description d'une expérience aléatoire et de l'introduction fréquentiste de la probabilité, on pourra reprendre l'article "Les outils mathématiques du calcul des probabilités", paru dans le bulletin de l'IREM de Besançon n° 47 (Mars 92).

(16) Pour un examen plus précis de l'articulation entre le programme de première et celui de terminale, on pourra consulter l'article "Remarques sur le programme de probabilités de terminale" paru dans le bulletin de l'IREM de Besançon n° 48 (Juin 92).

tout, l'apprentissage visé par le nouveau programme ne semble alors pas être effectif.

Devant cette situation, les professeurs de terminale n'ont pas le temps de reprendre l'enseignement des probabilités de première. Ils ont tendance à adopter un cours traditionnel basé sur les dénombrements et l'hypothèse d'une interprétation spontanée par les élèves des situations aléatoires présentées dans les exercices.

Cette observation mérite d'être nuancée par une enquête plus large que celle que j'ai pu effectuer. Le bilan de l'évaluation EVAPM 1 de 93 devrait compléter cette appréciation. En tout cas, elle souligne le fait que cet enseignement de nature fréquentiste n'est pas facile, loin s'en faut. Il laisse apparaître un certain malaise chez les enseignants qui, dans les meilleurs des cas, ont reçu une toute autre formation en probabilités. Ils font alors difficilement la part entre les descriptions des situations aléatoires et le modèle probabiliste.

Dans une autre terminale (une TC) les élèves semblaient avoir les idées plus claires sur la notion de probabilité, exprimée spontanément et exclusivement en termes de $\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$, avec l'hypothèse d'équiprobabilité explicitement recherchée.

Les élèves avaient alors tendance à poser l'équiprobabilité a priori sur n'importe quel exemple; on retrouve ici la démarche erronée de D'Alembert dans "croix ou pile" de l'encyclopédie.

Les outils de base de la combinatoire font l'objet d'un paragraphe particulier du programme de terminale, situé dans son texte avant la partie "probabilités". Il n'est pas évident qu'il faille organiser l'enseignement dans cet ordre-là : les permutations, arrangements et autres combinaisons n'interviendront pour résoudre certains problèmes simples de probabilités que dans les situations d'équiprobabilité des événements élémentaires. Introduits trop tôt, ils risquent d'être finalisés trop restrictivement aux probabilités, tout en limitant à celles-ci le champ d'exercices et de problèmes accessibles en terminale.

b) Variables aléatoires et lois

Ces deux notions, intimement liées, sont essentielles pour développer véritablement le calcul des probabilités au niveau de l'efficacité qu'il a acquis maintenant.

Dégagées par Paul Lévy et A.I. Kintchine, elles fondent véritablement la démarche scientifique en probabilités. Le modèle théorique (Ω, P) élémentaire est complété par l'introduction du concept de variable aléatoire :

$$X : (\Omega , P) \longrightarrow (R , P_x).$$

L'étude des diverses probabilités images P_x de P par X (définition formelle de la loi de X) est en fait à la base de l'enseignement universitaire du calcul des probabilités.

En terminale, Ω est un ensemble fini, et une variable aléatoire X représente un caractère statistique c dont les valeurs observables sont les r valeurs possibles pour X : les x_i , qui sont alors interprétées comme les images par X des éléments de Ω.

L'étude de lois classiques, comme la loi binômiale par exemple, est hors programme. La donnée de la loi de X est alors réduite à celle des probabilités élémentaires $p_i = P(X = x_i)$ pour toutes les valeurs x_i .

En évacuant l'activité éminemment probabiliste de reconnaissance de lois dans des situations discrètes concrètes typiques (Bernoulli, binômiale, hypergéométrique, géométrique, Poisson), le programme pose la question didactique de l'intérêt de la notion de loi, dès lors qu'elle se limite à la liste des probabilités élémentaires associées aux valeurs x_i de la v.a. X.

Ainsi l'estimation fréquentiste des paramètres déterminant la loi d'une v.a. judicieusement définie dans un contexte reconnu, disparait du champ accessible en terminale et retire une partie de la cohérence de la démarche entreprise. Celle-ci sera reprise dans les classes post-Bac : BTS, IUT ou Université, ce qui donnera à l'enseignement des probabilités son second souffle.

Pour représenter la loi d'une v.a., on peut :

— donner un tableau de valeurs associant les probabilités élémentaires p_i aux x_i,

— donner un moyen de calcul pour obtenir p_i à partir de la valeur de x_i,

— donner les valeurs des probabilités cumulées : $F(x_i) = P(X \leq x_i)$ pour chaque x_i, on retrouve alors $p_i = F(x_i) - F(x_{i-1})$, (où F est la fonction de répartition de la loi de X). Le lien avec les fréquences cumulées est alors simple.

Il serait bon d'user effectivement de ces trois possibilités, notamment de la dernière qui associe la fonction de répartition aux fréquences cumulées introduites lors de l'étude des séries statistiques : si les x_i sont les valeurs discrètes possibles pour un caractère quantitatif χ, elles déterminent un classement de la population sur laquelle χ est défini.

Après tirage au hasard d'un élément de cette population, l'identification de la probabilité $p_i = P(X = x_i)$ avec la proportion (ou fréquence) f_i de la classe déterminée par la valeur x_i fut à l'origine du concept de probabilité.

La progression se poursuit alors naturellement par la démarche indiquée dans l'encadré suivant :

La somme $\sum_{i=1}^{k} p_i = P(X \leq x_k)$ est associée à la somme $\sum_{i=1}^{k} f_i$, qui est la fréquence cumulée (croissante) des classes déterminées par les valeurs du caractère c, inférieures ou égales à x_k

La cohérence de l'approche fréquentiste veut qu'à la moyenne des valeurs prises par le caractère de la série statistique, on associe l'espérance mathématique de la variable. La moyenne des valeurs x_i du caractère c, pondérée par les fréquences f_i est : $\bar{x} = \sum_{i=1}^{r} f_i x_i$, d'où

$E(X) = \sum_{i=1}^{r} p_i x_i$. L'espérance de X est ainsi la moyenne des valeurs x_i pondérées par les probabilités p_i.

De même la variance et l'écart type de X seront associés aux variance statistique et écart type statistique du caractère c :

$$\mathrm{Var}\ c = \sum_{i=1}^{r} f_i\,(x_i - \bar{x})^2 \qquad \text{d'où} \qquad \mathrm{Var}\ X = \sum_{i=1}^{r} p_i\,(x_i - E[X])^2 = E[(X-E(X))]^2,$$

faisant apparaître que la variance de X est la moyenne pondérée par les p_i des carrés des écarts des valeurs de X à son espérance mathématique

c) Probabilités conditionnelles

On aborde ici une partie délicate du programme de terminale, car la notion de probabilité conditionnelle rencontre aussi des conceptions erronées [17] : comment interpréter le fait qu'une connaissance partielle sur un résultat d'une expérience aléatoire permet de modifier la probabilité que l'on va attribuer à un événement qui lui est lié ?

La probabilité d'un événement dépendrait-elle des informations obtenues par l'observateur ?

Cette conception erronée est fortement présente chez les élèves.

Il faut donc faire comprendre qu'une donnée nouvelle sur le déroulement d'une expérience aléatoire modifie non pas les probabilités des événements mais la *description* de cette expérience, et par conséquent le modèle utilisé pour représenter les résultats possibles : on va changer de référentiel et déterminer une nouvelle répartition de probabilités sur le nouveau référentiel.

En fait, une approche fréquentiste va clarifier cette situation : pour estimer une probabilité conditionnelle, on procède à une *autre* expérience aléatoire.

Supposons en effet qu'à l'issue de l'expérience primitive, on sache que l'événement B est réalisé.

Reproduisons cette expérience un grand nombre de fois, en éliminant les épreuves où B n'est pas réalisé (on peut du moins imaginer cette répétition).

Cette procédure met en œuvre une autre expérience aléatoire : celle qui consiste à ne retenir l'issue de la précédente, *que* si B est réalisé, sinon on recommence.

Dans cette nouvelle expérimentation, la probabilité d'un événement A est déterminée par la fréquence des A intervenant dans les épreuves retenues. Elle est donc différente de la probabilité a priori de A, estimée par la fréquence des A intervenant dans l'ensemble des épreuves reproduisant l'expérience primitive.

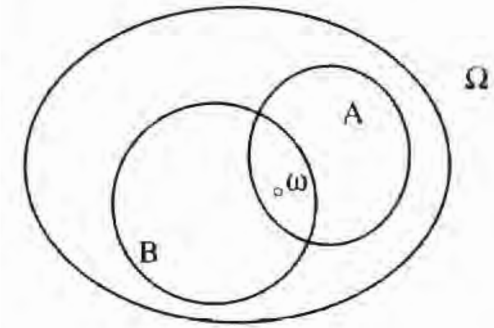
Une autre difficulté conceptuelle, décrite par Totohasina [35], est le lien naturel entre la notion de probabilité conditionnelle et celle de relation de cause à effet, dans le cadre d'épreuves successives. Dans la pratique, l'application des probabilités conditionnelles concerne le plus souvent des événements liés entre eux par une telle relation.

Cependant l'étude d'expériences compo-

(17) Les conceptions rencontrées chez les élèves et l'analyse d'un enseignement s'appuyant sur l'utilisation de shémas d'arbres fait l'objet de l'article d'André Totohasina : "L'introduction du concept de probabilité conditionnelle : avantages et inconvénients de l'arborescence", à paraître dans *Repères-IREM* n° 15

sées d'épreuves répétées n'est pas au programme (qui exclut la notion de probabilité produit) et il faudra viser juste dans les exercices pour s'en tenir à sa lettre.

La définition proposée par le programme de terminale se place résolument dans le cadre de la modélisation ensembliste, minimisant ainsi les obstacles dus à des conceptions subjectivistes. Reproduisons ici cette définition, afin de montrer comment elle peut prendre en compte les remarques précédentes :



Ω référentiel décrivant l'expérience aléatoire primitive,
B événement de probabilité non nulle P(B),
A événement dont on cherche la probabilité P(A),
ω événement élémentaire représentant l'issue de l'expérience.

Changement d'expérience : on sait que B est réalisé ($\omega \in B$). B est le nouveau référentiel. On désigne par P_B la probabilité des événements issus de cette expérience qui garantit la réalisation de B.

A ne peut alors être réalisé que par la réalisation de $A \cap B$. La probabilité de la réalisation de A, $P_B(A)$ sera donc celle de $A \cap B$ avec comme donnée: $P_B(B) = 1$.

Si les événements élémentaires sont équiprobables, ils le sont aussi bien dans Ω que dans B, et nécessairement :

$$P_B(A) = \frac{\text{cas favorables}}{\text{cas possibles}} = \frac{\text{card}(A \cap B)}{\text{card } B} = \frac{P(A \cap B)}{P(B)}$$

Mais si on est dans une situation de non-équiprobabilité, l'esprit fréquentiste du programme conduit à revenir à la définition de la probabilité par les fréquences, ainsi que cela est présenté dans l'encadré qui suit :

Désignons par $\mathcal{E}$ l'expérience primitive représentée par le référentiel Ω, et B un événement associé à $\mathcal{E}$ de probabilité non nulle : $P(B) \neq 0$.
Si l'on reproduit $\mathcal{E}$ un grand nombre n de fois, P(B) sera proche autant qu'on le veut de la valeur observée f_B de la fréquence d'apparitions de B :

$$f_B = \frac{\text{nombre d'occurrences des événements B}}{n}$$

Soit A un événement associé à $\mathcal{E}$, on s'intéresse à la réalisation conjointe de A et B. P(AiB) sera de même proche de :

$$f_{A \cap B} = \frac{\text{nombre de réalisations conjointes de A et B}}{n}$$

On s'intéresse à la fréquence d'apparition des A accompagnant la réalisation de B, celle-ci est donnée par :

$$f_{A/B} = \frac{\text{nombre de réalisations conjointes de A et B}}{\text{nombres d'occurences des événements de B}} \quad \text{d'où : } f_{A/B} = \frac{f_{A \cap B}}{f_B},$$

et la probabilité d'obtenir A, lorsque l'on considère l'expérience aléatoire $\mathcal{E}_B$ consistant à n'accepter un résultat de $\mathcal{E}$ que lorsque B est réalisé, est estimée par $f_{A/B}$.

La relation obtenue entre les fréquences conduit donc à la relation liant cette probabilité $P_B(A)$ aux probabilités a priori de $A \cap B$ et de B : $P_B(A) = \frac{P(A \cap B)}{P(B)}$, ce qui généralise cette même relation obtenue sous l'hypothèse d'équiprobabilité.

On remarque que $P_B(A)$ est proportionnelle à $P(A \cap B)$, ce qui est cohérent avec l'idée de réduire la probabilité de A à celle de $A \cap B$ lorsque B est réalisé, avec la contrainte : $P_B(B) = 1$.

On revient alors à la notation traditionnelle :

$P(A/B) = \frac{P(A \cap B)}{P(B)}$, ou encore $P(A \cap B) = P(A/B)P(B)$, quand $P(B) \neq 0$.

A partir du sens ainsi donné à la notion de probabilité conditionnelle, les élèves peuvent aisément relier la définition de l'indépendance stochastique de A et B, introduite dans le modèle sous la forme : $P(A \cap B) = P(A)P(B)$ avec l'idée concrète que l'on peut s'en faire ; A et B sont indépendants si la réalisation de B ne modifie pas *la valeur* de la probabilité de A d'un modèle à l'autre : $P(A/B) = P(A)$.

Cependant la notion d'indépendance reliée à des situations aléatoires concrètes se heurte encore à de nombreux obstacles ou conceptions erronées étudiés notamment dans les thèses de Sylvette Maury et d'André Totohasina. Les lecteurs intéressés pourront se reporter à ces travaux [32 et 35], ainsi qu'à l'article d'André Totohasina, à paraître dans *Repères-IREM* n° 15.

Le programme évite soigneusement de parler de causes d'un événement et exclut tout décalage dans le temps entre les réalisations de A et B. Cela permet de maintenir la symétrie entre A et B qui s'exprime dans la définition de la probabilité conditionnelle :

$P(A \cap B) = P(A/B)\ P(B) = P(B/A)\ P(A)$,
où $P(A)$ et $P(B)$ sont non nulles.

Et le programme exclut la formule de Bayes, formule très simple, mais qui pose des problèmes épistémologiques ardus comme nous l'avons vu : déterminer les probabilités des causes ayant observé leurs conséquences. En effet, comment interpréter dans une approche fréquentiste la notion de probabilité d'une cause ayant déjà produit, lors d'une expérience aléatoire, un résultat observé ? Dans quelle expérience aléatoire, même abstraite, cette cause serait-elle un événement ? Les conceptions bayésiennes ne se prêtent pas simplement à une intégration immédiate dans le processus de construction d'un modèle mathématique représentatif d'une expérience aléatoire et sont ici sources de difficultés didactiques certaines. Laissons provisoirement cette question ouverte et saluons la sagesse du programme qui renvoie cela aux classes post-bac.

En terminale, la réussite à des problèmes comme celui des jetons ci-dessous, permet de s'assurer de la compréhension et la maîtrise de la notion de probabilité conditionnelle, dès lors que cet outil est préféré aux dénombrements (dont l'utilisation est ici piégée) par des élèves qui reconnaîtront son efficacité.

De plus ce type de situation permet d'effectuer, en lui donnant un sens probabiliste, le produit de probabilités pour calculer celles d'événements composés, ce qui est une compétence essentielle en probabilités. Notons que l'exploitation (si l'on peut dire !) d'arbres schématisant des disjonctions composées est efficace dans ce type de problèmes d'enchaînement de probabilités [18]

Voici cet exercice test :

Dans un sac, il y a 3 jetons ; un avec deux faces noires, un avec deux faces blanches, un avec une face noire et une face blanche. On tire un jeton au hasard, il a une face blanche. Quelle est la probabilite que l'autre soit blanche ?

On peut certes résoudre cet exercice par dénombrement des cas. Mais les élèves ont du mal à modéliser l'expérience et confondent équiprobabilité des jetons et équiprobabilité des faces, hésitent à ordonner les faces et sont le plus souvent conduits à des résultats faux. En termes de probabilités conditionnelles, la situation s'éclaircit et le résultat s'impose :

Si B_1 est l'événement "la face regardée est blanche", et B_2 l'événement "l'autre face est blanche", on a :

$$P(B_2/B_1) = \frac{P(B_1 \cap B_2)}{P(B_1)}$$

Or : $P(B1 \cap B2) = 1/3$ (C'est le tirage du jeton blanc-blanc),

$P(B_1) = 1/2$ (C'est la proportion des faces blanches dans le jeu), d'où : $P(B_2/B_1) = 2/3$.

d) Résolution d'un paradoxe (le jeu de la voiture et des deux chèvres) [19]

Reprenons pour terminer, ce paradoxe qui, pour être résolu, suppose de recourir soigneusement aux connaissances acquises en première et terminale. Il peut être proposé aux élèves pour clore le chapitre des probabilités, car il montre de manière surprenante l'efficacité des outils construits. Rappelons ici succinctement la situation :

A la télévision américaine, il y a un jeu célèbre qui se déroule de la manière suivante:

— Le candidat se trouve devant trois portes fermées.

(18) Pour l'utilisation des arbres probabilistes, on pourra consulter l'excellent article de Bernard Parzysz, paru dans *Repères-IREM* n° 10 (Janvier 93).

(19) Ce paradoxe et ses enseignements ont fait l'objet de l'article "Paradoxes et lois de probabilités", paru dans *Repères-IREM* n° 13.

— Derrière les trois portes, on a disposé ***au hasard*** *deux chèvres et une voiture.*

— Le candidat désigne une porte de son choix qui reste fermée.

— L'animateur du jeu ouvre alors ***une autre porte*** *derrière laquelle il sait que se trouve une des deux chèvres.*

— Le candidat désigne à nouveau l'une des deux portes restant fermées; il gagnera l'objet, voiture ou chèvre, qui se trouve derrière.

Quelle probabilité a-t-il de gagner la voiture ?

Pour un élève de terminale, la bonne question est :

Comment formuler ce problème de probabilités ?

Car l'article cité montre que pour répondre à cette question de manière unique, il faut avoir précisé les conditions aléatoires du deuxième choix du candidat, c'est à dire avoir déterminé un modèle : ce choix s'effectue dans des conditions insuffisamment précises pour en imposer un parmi une infinité possible. D'où l'apparent paradoxe entre les deux raisonnements suivants:

a- *"Lorsque l'animateur ouvre la porte qui cache une chèvre, il n'apporte aucune information au candidat sur les places respectives de l'auto et de l'autre chèvre qui sont derrière les deux autres portes, puisque le règlement du jeu prévoit par avance cette action. Le candidat est alors ramené à un choix parfaitement hasardeux entre les deux portes qui lui restent, il a donc une chance sur deux de gagner la voiture."*

b- *"Le candidat, lors de son premier choix a une chance sur trois de choisir la porte de la voiture et deux chances sur trois de choisir la porte ouvrant sur une chèvre. Si, au deuxième choix, il désigne la même porte qu'au premier, quelle que soit la porte ouverte par l'animateur, sa probabilité d'avoir choisi la voiture reste 1/3. Si, par contre, il change de porte, "gagner la voiture" est alors le même événement que "avoir choisi une chèvre lors du premier choix". Il a donc deux chances sur trois de gagner la voiture".*

Troublant, non ? Lequel des raisonnements a-t-il votre faveur ?

Il est intéressant de remarquer que les professeurs-stagiaires (PLC2 Besançon, 1992) de mathématiques auxquels ce problème avait été formulé dans ces termes, ont tous choisi par écrit le raisonnement a), malgré un entraînement préalable aux exercices de probabilités.

Souvent, les personnes interrogées ne sont pas convaincues par le raisonnement b). Par contre, la probabilité 2/3 est acceptée quand on propose d'imaginer que ce jeu fait intervenir 1000 chèvres dont 999 seraient dévoilées par l'animateur !

Cela me suggère cette remarque d'ordre didactique : quand la solution d'un problème semble intuitivement évidente, un sujet, élève ou enseignant, n'a pas spontanément recours à des connaissances de type scientifique, même si celles-ci sont maîtrisées et disponibles, lorsque l'appel à une conception naïve, reposant sur les premiers apprentissages, semble suffire pour répondre à la question.

Notons ici qu'un argument de nature fréquentiste a plus de poids que l'élégante démonstration que l'on peut trouver dans *Repères-IREM* n° 13.

Revenons donc à une interprétation fréquentiste de cette situation.

Si le jeu ne devait avoir lieu qu'une fois, ce problème de calculer la probabilité de gagner la voiture n'aurait pas beaucoup de sens. En effet, une fois la deuxième porte ouverte, il n'y a plus de probabilité.

Mais là où la valeur de cette probabilité prend de l'importance, c'est quand le jeu est reproduit chaque semaine. Cette valeur intéresse alors bigrement le producteur de l'émission qui doit présenter un budget acceptable. Il doit donc avoir une idée au départ du nombre des candidats qui vont gagner la voiture.

Mais ce producteur a lu *Ars conjectandi* de Bernoulli et nous invite à une relecture :

"qui aurait sur la nature de l'esprit humain, ou sur l'admirable fabrique de notre corps une vue suffisante pour oser déterminer dans les jeux, qui dépendent en totalité ou en partie de la finesse de celui-là ou de l'agilité de celui-ci, les cas qui peuvent donner la victoire ou l'échec à celui-ci ou à celui-là des concurrents ?". [5, p. 42].

Le producteur sait que pour le calcul du coût de l'émission, il doit estimer les probabilités que les candidats adoptent telle ou telle stratégie. Comme l'indique Bernoulli, leur choix de stratégie reste hasardeux et échappe à l'approche pascalienne :

" car des faits dépendent de causes tout à fait cachées, et destinées de plus par l'indénombrable variété des assemblages à se jouer éternellement de notre recherche". [ibid]

Bernoulli indique alors (évoquant cet autre chemin dans la citation rapportée au début de la deuxième partie) qu'il ne reste que l'estimation par les fréquences observées des comportements des candidats au cours d'un grand nombre d'émissions.

e) Perspectives

Ainsi, l'approche fréquentiste de la notion de probabilité éclaire d'un nouveau jour ce calcul des probabilités jugé encore rébarbatif. Les élèves y découvrent un lien réel avec des situations concrètes, tout en comprenant mieux le rôle des mathématiques.

L'enseignement actuel des probabilités a franchi une étape déterminante en quittant le terrain de la théorie formelle (espaces probabilisés, tribus, mesures...), telle qu'on la retrouve dans l'enseignement universitaire (impropre par conséquent à éclairer les enseignants du second degré sur les enjeux de leurs programmes), et en abandonnant la conception très réductrice des probabilités comme champ d'applications de la combinatoire.

A partir d'un point de vue fréquentiste, les programmes de 1991 fixent donc comme véritable objectif de l'enseignement des probabilités, l'acquisition d'un outil de résolution de problèmes réels.

Mais ils ne sont qu'au milieu du gué. En laissant ouverte la confusion entre fréquence et probabilité (on peut relire certains manuels), ils semblent trancher en faveur d'une position épistémologique : la probabilité serait une fréquence limite (et la théorie serait bâtie sur cette assimilation, telle que Von Mises a tenté de la fonder au début du vingtième siècle).

Encore un pas à franchir pour donner à la probabilité le statut de concept mathématique, abstraction d'une notion concrète, calculable à partir d'hypothèses heuristiques posées a priori ou estimable à partir de l'observation de fréquences.

La clarification de cette démarche passe par le choix didactique de séparer description du réel et modèle mathématique. Le programme conçu pour la série L semblait s'y engager, jusqu'à ce que son application soit remise en cause par des décisions récentes ; qu'en sera-t-il de l'avenir de l'enseignement des probabilités ?

Quoi qu'il en soit, cet avenir dépend avant tout de la formation des enseignants, donc, pour une part, du travail qui se fait dans les IREM et qui s'exprime notamment dans *Repères-IREM*.

En comprenant mieux le lien entre statistiques et probabilités, nos collègues découvrent dans le calcul des probabilités un outil puissant pour les applications. Puissent-ils afficher sur la porte de leur salle de classe :

"nul n'entre ici s'il n'est probabiliste" !

RÉFÉRENCES BIBLIOGRAPHIQUES DES OUVRAGES CITÉS

Textes historiques (par ordre d'entrée en scène)

[1] PLATON (427-347 av JC) : *La République*, texte grec (380 av JC) et traduction E. Chambry - Collection Budé, Les Belles Lettres, Paris -1932.

[2] ARISTOTE (384-322 av JC) : *Physique*, texte grec et traduction H. Carteron - Les Belles Lettres, Paris - 1969 -*Organon, les seconds analytiques*, traduction J. Tricot, Paris, Vrin - 1979.

[3] CARDAN Jérôme (1501-1576) : *Liber de ludo aleae* - 1526, publié à Lyon en 1663.

[4] PASCAL Blaise (1623-1662) et FERMAT Pierre (1601-1665) : *Correspondance entre Pascal et Fermat* -1654, *in* Œuvres de Pascal, ed. J. Chevalier, La Pléïade, Paris, Gallimard 1963.

[5] BERNOULLI Jacques (1667-1748) : *Ars conjectandi* (1713), traduction de N. Meusnier - publication de l'IREM de Rouen - 1987.

[6] BAYES Thomas (1702-1761) : *Essai en vue de résoudre un problème de la doctrine des chances* - traduction de J.-P. Cléro, *Cahiers d'Histoire et de Philosophie des Sciences*, n° 18, 1987.

[7] D'ALEMBERT Jean le Rond (1717-1783) : *La grande Encyclopédie*, env.1750-1780, et correspondance.

[8] CONDORCET Antoine (1743-1794) : *Elémens du calcul des probabilités*, (publié en 1805) - publication de l'IREM de Paris VII - 1986.

[9] LAPLACE Pierre-Simon (1749-1827) : *Essai philosophique sur les probabilités*, 1825 -Ed. Christian Bourgeois - 1986.

— [9'] : -*Théorie analytique des probabilités* - Courcier imprimeur Paris - 1814.

[10] COURNOT Antoine Augustin (1801-1877) : *Exposition de la théorie des chances et des probabilités- in* œuvres complètes, vol. 1, ed. Vrin, 1984.

[11] BERTRAND Joseph (1822-1900) : *Calcul des probabilités* - Gauthier-Villars - Paris - 1889.

[12] POINCARÉ Henri (1854-1912) : *Calcul des probabilités* -Jacques Gabay -1987 (Réédition de l'édition Gauthier-Villars de 1912).

— [12'] : *La science et l'hypothèse* - Flammarion-1902, réédité par Champs/Flammarion - 1968.

[13] LEBESGUE Henri (1875-1949) : *Leçons sur l'intégration* - 1904 - réédité par Gabay-1987.

[14] BOREL Emile (1871-1956) : *Leçons sur la théorie des fonctions* - Gauthier-Villars - Paris - 1898.

— [14'] : *Eléments de la théorie des probabilités* - Hermann et fils - 1909.

— [14"] : *Le hasard* - 1914. Réedition P.U.F. 1943.

[15] KOLMOGOROV Andréi (1903-1987) : *Grundbegriffe der Wahrscheinlichkeitsrechnung* - Springer-1933

Histoire des probabilités et de la statistique

[16] COMMISSION INTER-IREM HISTOIRE ET ÉPISTEMOLOGIE : *Mathématiques au fil des âges*- Gauthiers-Villars - 1987

[17] BRU Bernard : "Petite histoire du calcul des probabilités" - Article paru dans *Fragments d'histoire des mathématiques*- Brochure APMEP n° 41-1981.

[18] DROESBEKE Jean-Jacques et TASSI Philippe : *Histoire de la statistique* - P.U.F. - 1990

Epistémologie

[19] BACHELARD Gaston : *La formation de l'esprit scientifique* -1938- Librairie philo, Vrin, 1965 (4e edition).

[20] FRECHET Maurice : *Les mathématiques et le concret* - 1955-Coll. philosophie de la matière, PUF

[21] DHOMBRES Jean : *Nombre, mesure et continu. Epistémologie et histoire*- Publication de l'IREM de Nantes, CEDIC/Nathan, 1978.

[22] EKELAND Ivar : *Au hasard* - Seuil - 1991.

[23] NOËL Emile et France-Culture : *Le hasard aujourd'hui* - Points/Sciences -1991.

[24] PRIGOGINE Ilya et STEENGERS Isabelle : *La nouvelle alliance* - Folio/Essais, Gallimard, 1979.

[25] RUELLE David : *Hasard et chaos* - Ed. Odile Jacob - 1991.

Manuels d'enseignement

[26] CALOT Gérard : *Cours de calcul des probabilités* - Dunod décision - 1967.

[27] ENGEL Arthur : *Les certitudes du hasard* -Aleas Lyon - 1990.

[28] RENYI Alfred : *Calcul des probabilités* - Dunod - 1966.

[29] SAPORTA Gilbert : *Probabilités, analyse des données et statistiques* - Technip - 1992

[30] VENTSEL Hélène : *Théorie des probabilités* - Ed. Mir, Moscou - 1973.

Didactique des probabilités

[31] BROUSSEAU Guy : "Généralités sur l'enseignement des probabilités au niveau élémentaire" - *in Actes de la 26e rencontre de la CIEAEM*,- IREM de Bordeaux, Août 1974, pp. 66-123.

[32] MAURY Sylvette : *Contribution à l'étude didactique de quelques notions de probabilité et de combinatoire à travers la résolution de problèmes* -Thèse d'Etat, Montpellier I-1986.

[33] ZAKI Moncef : *Traitements de problèmes de probabilités en situation de simulation*, Thèse de doctorat , IRMA-ULP Strasbourg -1990.

[34] BORDIER Jacques : *Un modèle didactique utilisant la simulation sur ordinateur, pour l'enseignement de la probabilité* -Thèse de doctorat, Paris VII- 1991.

[35] TOTOHASINA André : *Méthode implicative en analyse de données et application à l'analyse de conceptions d'étudiants sur la notion de probabilité conditionnelle*, Thèse de doctorat, IRMAR Rennes I -1992.