Galilée ou Descartes ? Étude d'un scénario d'introduction historique au calcul des probabilités

Éric Butz

GALILÉE OU DESCARTES ?
Étude d'un scénario d'introduction historique au calcul des probabilités

Éric BUTZ
Lycée Lislet-Geoffroy, Saint-Denis.

RÉSUMÉ. – Cette étude prospective propose des activités qui s'appuient sur l'étude de textes anciens, scientifiques et philosophiques. En rapport étroit avec les programmes scolaires, elle veut participer à la formation continuée de tous les enseignants. En effet, les nouveaux programmes de mathématiques, de la Seconde à la Terminale, imposent une nouvelle approche de l'enseignement des statistiques et des probabilités. Ils induisent, d'une part, des études historiques et épistémologiques, et, d'autre part, rendent incontournable l'utilisation de simulations, de modélisations et de logiciels informatiques. Dans ce qui suit, nous présentons un scénario permettant la construction de savoirs mathématiques associés au calcul des probabilités, à travers une initiation à la recherche scientifique.

ABSTRACT. – *This prospective article suggests activities based on the study of ancient texts on science and philosophy. Closely related to educational syllabuses it aims to participate in the continuous training of teachers. Indeed, the new syllabuses in maths, from seconde to terminale forms, have imposed a new approach to statistics and probabilities. Besides including historical and epistemological studies, they make it necessary to use simulations, modelisations and teaching software. In the following article we will present a pattern meant for the acquisition of mathematical knowledge associated with probabilities through an initiation to scientific research.*

Les nouveaux programmes de mathématiques du lycée imposent l'apprentissage du hasard (statistiques et probabilités) à l'aide d'expériences aléatoires, de simulations et de modélisations. Or la méthode expérimentale semble être étrangère aux pratiques actuelles de l'enseignant de mathématiques, alors qu'elle caractérise, dans les esprits, les autres sciences. Nous tenterons d'analyser cette situation à travers l'enseignement des statistiques et des probabilités, et de proposer une piste d'action possible[1].

1. Analyse

1.1. L'évolution des programmes

Dans les nouveaux programmes, l'introduction des probabilités à partir des statistiques constitue un véritable bouleversement par rapport aux référentiels précédents. Ainsi, en mai 1971, le programme des Premières CDE introduisait l'axiomatique du calcul des probabilités en s'appuyant sur un triplet de la forme (E, P(E), *p*). Le programme de la classe de Terminale C comportait alors quatre paragraphes dont les titres étaient, dans l'ordre :

1. Espaces probabilisés finis (E, B(E), *p*).
2. Espérance mathématique d'une variable aléatoire à valeurs dans $\mathbb{R}$.
3. Inégalité de Bienaymé-Tchebychev.
4. Description statistique d'une population ou d'un échantillon.

C'était l'omniprésence de l'axiomatique de Kolmogorov sur les algèbres de Boole ou les tribus de Borel. L'enseignement des statistiques se traitait après celui des probabilités. Dans un manuel de cette époque, on pouvait lire : « L'étude d'un problème de probabilités commence nécessairement par la détermination de l'algèbre des événements B(E), puis par celle de la probabilité *p*. La détermination de la probabilité *p* conduit fréquemment à des problèmes de dénombrements » (Gautier *et al*., 1975, p. 357).

En 1986, les algèbres et les tribus ont disparu, remplacées par un « univers » des événements, mais l'ordre combinatoire-probabilités est inchangé.

En 1992, on lit dans le programme : « Combinatoire-probabilités ». Là encore, l'ordre est imposé, mais « toute théorie formalisée est exclue ». Les règles de calculs sont « déduites » de représentations des situations sous formes de tableaux ou d'arbres.

Même esprit dans les programmes des classes de Première et Terminale S de juin 1997. Les règles de calcul concernant les probabilités des événements « A ou B » et « A et B » sont associées à la transcription d'une expérience aléatoire sous forme d'arbre ou de tableau. La probabilité d'un chemin est le produit des probabilités rencontrées sur les branches. La probabilité d'un événement est la somme des probabilités des chemins correspondant à cet événement. La notion d'espérance est associée à celle de variable aléatoire.

Historiquement, les règles de calcul des probabilités, en particulier celle qui concerne *p*(A ou B), ont été introduites à partir de la notion d'espérance et du calcul des dénombrements finis, puis elles ont été généralisées, par analogie, aux cas infinis et continus.

Dans le programme du 31 août 2000, après le paragraphe sur les statistiques, on lit : « Probabilités. Définition d'une loi de probabilité sur un ensemble fini. Espérance, variance, écart-type d'une loi de probabilité. Probabilité d'un événement, de la réunion et de l'intersection d'événements. Cas de l'équiprobabilité. [...] Modélisation d'expériences aléatoires de référence ». Et, dans la colonne des modalités de mise en œuvre, on trouve : « Le lien entre loi de probabilité et distributions de fréquences sera éclairé par un énoncé vulgarisé de la loi des grands nombres. »

Dans le document d'accompagnement, on veut nous apprendre à faire jongler nos élèves avec les lois et leurs simulations pour pouvoir les appliquer, en biologie par exemple. Mais la seule référence au calcul de p(A ou B) est : « On évitera tout développement théorique sur le langage des événements et le calcul ensembliste qui en découle : ces notions et la pratique de la logique qu'ils impliquent (étude du complémentaire de l'événement A ou B, ou de l'événement A et B) s'acquièrent au fil d'exercices. » Nos élèves devront être à la fois Galilée pour les expériences et Descartes pour la connaissance innée des règles du calcul des probabilités !

Un tel écart entre le niveau demandé et la réalité sur le terrain fait que, actuellement, par manque de temps et de matériel, on ne peut présenter en classe que des simulations préparées entièrement par l'enseignant, sans aucune action de la part des élèves. Cela revient à cantonner l'élève dans une utilisation passive de l'ordinateur, alors qu'il faudrait commencer par une véritable introduction aux méthodes de recherche scientifique.

1.2. Nos hypothèses

Je fais les mêmes hypothèses que le groupe technique disciplinaire qui publie les programmes (GTD, devenu depuis peu GEPS) : hypothèses selon lesquelles les élèves ont acquis le niveau nécessaire et disposent du temps et du matériel souhaités. Le scénario ci-après s'inscrit dans une progression sur trois ans et suit les programmes des trois classes de lycée. Le découpage par niveaux relèverait d'une autre étude.

Il semble que les scientifiques soient d'accord pour définir l'action de modéliser comme le passage d'une série de mesures obtenues expérimentalement à un modèle abstrait. Au lycée, on peut supposer que la démarche se résume, dans un premier temps, à trouver une relation fonctionnelle associée à un nuage de points. Les fonctions sont choisies empiriquement dans un catalogue de fonctions connues. Puis, le degré d'adéquation entre le modèle proposé et la réalité est estimé par d'autres mesures. Enfin, une décision est prise : si le modèle est confirmé par ces nouvelles mesures, même approximativement dans les limites d'un cahier des charges, il pourra alors être utilisé pour des prévisions et des simulations approchées (dans ce premier cas, le scientifique doit toujours être prêt à améliorer un modèle retenu) ; dans le cas contraire, le modèle est rejeté et il faut en inventer un autre.

Ainsi, un écueil apparaît en classe : comment estimer la précision des différents modèles obtenus ? Devrons-nous et pourrons-nous les départager ?

Aucune méthode d'ajustement n'est indiquée dans les programmes. Devons-nous en exposer les principales ?

Très souvent, les scientifiques utilisent la méthode des moindres carrés pour ajuster une série de mesures par une fonction. Cette méthode, ignorée des programmes de la série S, est présente dans ceux de la série ES. Elle est implicite lorsqu'on utilise un tableur pour obtenir les courbes de tendance.

Dans le cas où la recherche d'une relation fonctionnelle échouerait, une autre démarche scientifique consiste à construire un modèle permettant d'obtenir de proche en proche de nouvelles valeurs. C'est le cas, par exemple, de la méthode d'Euler qui consiste à construire par points une solution approchée de l'équation différentielle $y' = f(t, y)$ avec la condition initiale $y(t_0) = y_0$.

1.3. Confrontation des programmes à nos hypothèses

Il ne semble pas que les buts affichés par les programmes soient en conformité, ni avec la volonté politique et sociale d'ouvrir les classes scientifiques au plus grand nombre, ni avec la réduction drastique des horaires d'enseignement. En particulier, comment accepter que le programme officiel affirme que « l'utilisation de logiciels requiert des connaissances et des compétences mathématiques que cette utilisation contribue en retour à développer », puis « insiste pour que cet aspect du lien entre mathématiques et informatique soit travaillé à tous les niveaux », et se contredise enfin par ces mots : « compte tenu de l'horaire imparti et des débats en cours, il n'est proposé aucun chapitre d'informatique » ?

Sans heures dédoublées pour un apprentissage sérieux des logiciels, ce programme accentuera la différence entre les élèves qui peuvent avoir un ordinateur à la maison et les autres, entre les lycées « riches » et les autres. Ainsi, à mon avis, la référence à l'utilisation de l'informatique et des logiciels relève d'un vœu pieux, et c'est bien dommage. Une telle réforme ne peut pas se faire à moyens constants.

1.4. L'enseignement des mathématiques depuis 1968

À partir de 1968, formés à l'école de Bourbaki, nous nous étions habitués, après présentation de quelques exemples construits spécialement pour cela, à enseigner chaque nouvelle notion de mathématiques à partir d'axiomes et de définitions ; des règles de logique nous permettaient alors de « construire la mathématique ». La pensée de Descartes régnait en maître : « Ne rien emprunter aux représentations sensibles ; tout demander à l'entendement ; négliger tous les résultats de l'observation et considérer seulement ces semences de vérité qui sont naturellement en nous » (Namer, 1979, p. 16).

Les notions premières de mathématiques étaient réputées innées. Mais si, dans l'école républicaine, on cachait consciencieusement le Dieu évoqué très souvent par Descartes pour avoir des idées, combien d'élèves ont dû faire un acte de foi pour admettre les axiomes donnés ou imposés par le maître ?

Revirement de la situation en l'an 2001 : les nouveaux programmes de Terminale S parlent des « sciences mathématiques », ce qui témoigne peut-être de la fin du clivage entre mathématiques pures et mathématiques appliquées.

2. Descartes ou Galilée ?

Depuis Aristote (384-322 av. J.-C.), seule compte la démonstration logique. Les lois sont posées *a priori* dans le seul but de « sauver les apparences ». L'expérience n'est réalisée qu'ultérieurement, pour vérifier le modèle et le justifier, non pour le créer. C'est ainsi que, pendant presque deux millénaires, les principaux modèles qui décrivent les mouvements des planètes combinent des mouvements circulaires et uniformes, de centre la Terre, et sont tous contraires à la réalité.

Galilée (1564-1642) reprend les travaux et les méthodes d'Archimède et utilise les mathématiques dans le but d'expliquer les lois de la nature :

> « Ainsi, à 23 ans, en 1586, il énonçait déjà les conditions d'une connaissance de type nouveau, indépendante de tout dogmatisme, qu'il fût métaphysique ou mathématique. Une simplification opérée par l'esprit devait déterminer, à titre d'hypothèse méthodique, le choix des facteurs essentiels de la réalité en question, et le contrôle des éléments de l'expérience. Le savant devait inventer le dispositif expérimental susceptible de vérifier son hypothèse, afin de s'assurer, dans la mesure du possible, qu'aucune influence étrangère n'intervenait dans l'exécution de son expérience. Enfin, il devait exprimer mathématiquement les rapports entre les faits examinés, éventuellement en déduire les conséquences mathématiques, qui, à leur tour, seraient expérimentalement vérifiées. Ainsi, pour Galilée, les mathématiques abstraites ne se substituaient pas à la réalité, comme le fera Descartes, quelque trente ans plus tard, quand il aura reçu du ciel la révélation surnaturelle d'une mathématique universelle » (Namer, 1979, p. 58).

La méthode de Galilée impose des allers et retours entre l'observation et la mathématisation : une loi (aujourd'hui on parlerait plutôt de modèle) n'est jamais définitive, elle pourra être soit abandonnée, soit améliorée en cas de nouvelle découverte.

Au contraire, Descartes était « persuadé que seule la révélation mathématique et transcendante dépouillée de toute contamination matérielle, contenait la vérité absolue » (Namer, 1979, p. 58). Le dilemme est là :

> « Les mathématiques tendaient à devenir chez Galilée des techniques, lui permettant de poser des questions à la nature, et d'en recevoir des réponses neuves. Les mathématiques n'étaient que l'indispensable instrument de la recherche et de l'organisation des faits. Dans la perspective cartésienne, les mathématiques constituent le langage par lequel l'être de tout exprime son essence. Elles ne sont pas l'instrument de la recherche, mais la lumière grâce à laquelle l'intelligence humaine peut accéder à la contemplation de l'ordre cosmique » (*Ibid.*, p. 40).

On peut ainsi entrevoir les difficultés de l'enseignement des probabilités. Nous pensons que le hasard dépend plus de la « philosophie naturelle » que de la géométrie ou de l'arithmétique. Par exemple, le fait d'avoir démontré que la probabilité d'obtenir une face lors d'un lancer de dé est 1/6 ne vous donne aucune indication sur la face qui apparaîtra lors d'un prochain lancer. Ce n'est pas parce que l'as n'est pas sorti lors de cinq lancers consécutifs qu'il va sortir au sixième coup. Dans cette partie de notre enseignement, nous sommes bien à l'école de Galilée où tout est provisoire. Ce doute est en contradiction avec les certitudes des enseignants de mathématiques qui sont plus habitués à travailler sur la « Vérité » mathématique immuable (à chaque fois que l'on refait le même exercice, on est certain d'obtenir le même résultat).

Au contraire, faire des probabilités c'est, en quelque sorte, mesurer notre ignorance, comme le souligne Laplace lorsqu'il affirme :

> « Nous attribuons les phénomènes qui nous paraissent arriver et se succéder sans aucun ordre, à des causes variables et cachées, dont l'action a été désignée par le mot hasard, mot qui n'est au fond que l'expression de notre ignorance. La probabilité est relative, en partie, à cette ignorance, et en partie, à nos connaissances » (Laplace, 1812, p. 177-178).

L'utilisation de l'outil informatique n'est-elle pas en contradiction avec la simulation du hasard et doit-on, comme Descartes, rejeter « toute contamination matérielle » ? Comment passer de données empiriques, qualitatives et subjectives à des quantités mesurables et objectives ?

3. Évolution des idées concernant le point de départ du calcul des probabilités

Les rappels suivants sont synthétiques et fragmentaires, donc critiquables. Les définitions et lois de probabilités sont rarement l'œuvre d'un seul mathématicien, mais plutôt l'aboutissement de conjectures qui évoluent au fur et à mesure des nombreuses lettres échangées entre les différents chercheurs.

3.1. Blaise Pascal (1623-1662)

Pascal veut géométriser le hasard à partir du problème des partis. Pour cela, il utilise principalement la combinatoire (1654). Ses travaux sont basés sur des cas particuliers numériques. Il n'essaie pas d'utiliser des épreuves répétées pour prévoir l'avenir.

3.2. Jacques Bernoulli (1654-1705)

Dans *Ars conjectandi* (1713), Bernoulli démontre, à partir des combinaisons, la loi faible des grands nombres, c'est-à-dire, pour ne prendre qu'un exemple, que la fréquence d'apparition d'une face d'un dé converge vers 1/6 lorsque le nombre de lancers tend vers l'infini.

3.3. Abraham de Moivre (1667-1754)

Dans la préface de son traité *The Doctrine of Chances* (1718), Moivre annonce qu'il veut « donner une méthode de calcul des effets du hasard [...] et par ce moyen fixer certaines règles pour estimer jusqu'à quel point plusieurs espèces d'événements peuvent plutôt relever d'un dessein que du hasard » (cité dans *Cahiers d'histoire et de philosophie des sciences*, n° 18, 1988, p. 149). Dans ce but, il introduit en probabilités les méthodes de l'analyse et énonce la formule des probabilités composées : $p(\text{A et B}) = p(\text{A}).p(\text{B/A})$ (probabilité conditionnelle de B par rapport à A).

3.4. Thomas Bayes (1702-1761)

Dans son *Essai en vue de résoudre un problème de la doctrine des chances* (1763), Bayes est l'un des premiers à formaliser les règles du calcul des probabilités. En particulier, il définit la probabilité à partir de l'espérance, et il démontre la formule permettant le calcul de la probabilité de la réunion de deux événements incompatibles à partir de la linéarité de l'espérance (voir annexe 1). Plus tard, ce résultat sera pris pour axiome par Kolmogorov (voir annexe 2).

3.5. Pierre-Simon Laplace (1749-1827)

En 1812, dans sa *Théorie analytique des probabilités*[2], Laplace reprend les travaux de Moivre et Bernoulli. Il applique les résultats de l'analyse aux probabilités (théorème central limite, méthode des moindres carrés). Il justifie, en particulier, la méthode scientifique qui consiste à appliquer le « calcul des probabilités, à la recherche des phénomènes et de leurs causes » (1812, p. 363). C'est le début des statistiques au sens actuel. Par ses travaux, Laplace justifie l'utilisation de résultats du passé, qu'ils soient obtenus par observations ou par expériences, pour expliquer des phénomènes et prévoir leurs évolutions :

> « Je m'attache surtout, à déterminer la probabilité des causes et des résultats indiqués par les événements considérés en grand nombre, et à chercher les lois suivant lesquelles cette probabilité approche de ses limites, à mesure que les événements se multiplient. Cette recherche mérite l'attention des géomètres, par l'analyse qu'elle exige : c'est là principalement que la théorie de l'approximation des formules fonctions de grands nombres, trouve ses applications les plus importantes. Cette recherche intéresse les observateurs, en leur indiquant les milieux qu'ils doivent choisir entre les résultats de leurs observations, et la probabilité des erreurs qu'ils ont encore à craindre. Enfin, elle mérite l'attention des philosophes, en faisant voir comment la régularité finit par s'établir dans les choses même qui nous paraissent entièrement livrées au hasard, et en dévoilant les causes cachées, mais constantes, dont cette régularité dépend » (*Ibid*., p. 2).

3.6. Antoine Augustin Cournot (1801-1877)

En 1851, dans un *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*[3], le philosophe et mathématicien Cournot, spécialiste de probabilités et d'économie, définit d'abord « des séries indépendantes, c'est-à-dire qui se développent parallèlement ou consécutivement, sans avoir les unes sur les autres la moindre influence, ou (ce qui reviendrait au même pour nous) sans exercer les unes sur les autres une influence qui puisse se manifester par des effets appréciables » (Cournot, 1851, p. 38).

Ce qui lui permet d'expliquer ce qu'est le hasard : « 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 événements fortuits, ou des résultats du hasard » (*Ibid*., p. 39).

Puis la probabilité : « Mais on a donné le nom de probabilité mathématique à la fraction qui exprime le rapport entre le nombre des chances favorables à un événement et le nombre total des chances » (*Ibid*., p. 45).

Cournot exprime aussi l'idée contenue dans les nouveaux programmes et qu'il faut faire assimiler aux élèves de Seconde : « Dans une épreuve aléatoire, si l'on répète un très grand nombre de fois la même épreuve, le rapport entre le nombre des épreuves qui amènent l'événement *a* et le nombre total des épreuves doit différer très peu de la probabilité de l'événement *a* » (*Ibid.*, p. 46).

Mais il va plus loin puisqu'il nous expose une méthode pour les sciences mathématiques : la « probabilité philosophique », associée à l'induction et à l'analogie, permet de découvrir des lois à partir d'un petit nombre d'observations ou mesures consécutives. En effet, « les algébristes n'ont pas de peine à démontrer qu'on peut toujours assigner une loi mathématique, et même une infinité de lois mathématiques différentes les unes des autres, qui lient entre elles les valeurs successivement amenées » (*Ibid.*, p. 54).

Mais alors comment faire le bon choix ? Tout simplement en observant

> « la simplicité d'une loi, car si pourtant la loi mathématique à laquelle il faut recourir pour lier entre eux les nombres observés était d'une expression de plus en plus compliquée, il deviendrait de moins en moins probable, en l'absence de tout autre indice, que la succession de ces nombres n'est pas l'effet du hasard » (*Ibid.*, p. 54).

> « Lorsqu'à l'inspection d'une suite de valeurs numériques obtenues ainsi qu'il a été expliqué plus haut, on a choisi, entre l'infinité de lois mathématiques susceptibles de les relier, celle qui nous frappe d'abord par sa simplicité, et qu'ensuite des observations ultérieures amènent d'autres valeurs soumises à la même loi, la probabilité que cette marche régulière des observations n'est pas l'effet du hasard va évidemment en croissant avec le nombre des observations nouvelles : elle peut devenir et même elle devient bientôt telle qu'il ne reste plus à cet égard le moindre doute à tout esprit raisonnable. Si au contraire la loi présumée ne se soutient pas dans les résultats des observations nouvelles, il faudra bien l'abandonner pour la suite et reconnaître qu'elle ne gouverne pas l'ensemble de la série ; mais il ne résultera pas de là nécessairement que la régularité affectée par les observations précédentes soit l'effet d'un pur hasard ; car on conçoit très bien que des causes constantes et régulières agissent pour une portion de la série et non pour le surplus » (*Ibid.*, p. 55).

Cette méthode est explicitée au paragraphe 5.3.1 à travers la découverte de la troisième loi de Kepler.

3.7. Andreï Kolmogorov (1903-1987)

En 1933, le russe Kolmogorov présente une théorie axiomatique des probabilités à partir d'axiomes simples définissant une probabilité (voir annexe 2). Elle permet les allers et retours entre les fréquences obtenues comme résultats d'expériences répétées un grand nombre de fois et un modèle abstrait :

> « 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 » (Kolmogorov, cité par Bordier (1991) et par Henry (1999)).

4. Problèmes liés à l'usage de l'ordinateur

Les contemporains de Galilée lui reprochaient déjà d'utiliser une lunette pour ses découvertes, ce qui pouvait, disaient-ils, ne pas donner une image exacte de la réalité. L'utilisation de cet instrument ne modifiait-elle pas l'image observée ? Ce problème est récurrent. La construction d'un point d'une figure géométrique à l'aide de droites et de cercles est-elle plus juste que celle obtenue à l'aide de courbes « mécaniques » ? Et si la construction est impossible à la règle et au compas, faut-il s'interdire le recours à d'autres instruments ? La duplication du cube et la trisection de l'angle sont des exemples célèbres de problèmes de constructions impossibles à la règle et au compas qui ont nécessité l'emploi d'autres instruments.

Alors, pourquoi tant de réticences à utiliser un ordinateur pour simuler le hasard ? Il est vrai qu'il est difficile d'admettre qu'un algorithme déterministe puisse nous aider à simuler des tirages de nombres au hasard. De très nombreux tests mathématiques permettent de vérifier expérimentalement que les nombres pseudo-aléatoires ainsi obtenus permettent de simuler le hasard avec une assez bonne approximation, compatible avec nos connaissances actuelles. Ne tombons pas dans le piège de penser que les résultats des lancers d'un vrai dé sont obtenus davantage « au hasard » que ceux d'une simulation sur ordinateur. En effet, la notion de dé parfaitement équilibré n'est qu'une vue de l'esprit, uniquement réalisable comme hypothèse d'un exercice de mathématiques, mais ce n'est certainement pas la qualité d'un dé véritable. Et si, un jour, un mathématicien de génie améliorait ces fonctions aléatoires, soyons prêts, comme Galilée, à reprendre nos travaux pour améliorer la précision de nos simulations : c'est tout à l'honneur de l'esprit scientifique de pouvoir se remettre en cause. On est très loin des certitudes éternelles de Descartes.

5. Une progression faisant intervenir l'épistémologie et l'histoire des mathématiques, conformément au programme

5.1. Quelques remarques attendues de la part des élèves

Vérifier par simulation que, pour un dé, $p(a) = 1/6$, est-il constructeur de sens ? Il est légitime de se poser la question, car tous les enseignants ont entendu leurs élèves s'exprimer par des phrases du type : « Mais à quoi ça sert de faire des simulations pour trouver que j'ai une chance sur six de prévoir la sortie du numéro 2 lorsque je jette un dé ? Je le sais déjà » ; « Ça ne sert à rien de simuler le hasard ; de toute manière on n'est jamais certain du résultat : ce n'est pas une simulation qui va me faire gagner au loto ! »

5.2. Une initiation à la démarche scientifique

Une véritable démarche scientifique impose, au contraire, la recherche d'une loi inconnue et non triviale. C'est alors qu'elle permettra la construction du sens par les élèves. C'est pour cela que nous suivrons l'évolution historique, en commençant par un point qui nous paraît oublié par les programmes, bien qu'il soit important : la recherche d'une relation entre deux variables, dans le but d'utiliser l'approximation d'un nuage de points par une fonction (cette première recherche n'étant pas liée au hasard).

Pour cela, nous utiliserons des extraits de textes historiques et l'outil informatique, essentiellement les tableurs. Notre démarche suppose que les élèves ont des connaissances sur :

1) la notion de fonction ;

2) l'idée d'essayer de remplacer un nuage de points discrets par une courbe continue dont on cherchera une équation ;

3) la manière d'obtenir les courbes de tendance et leurs équations sur un tableur (éventuellement, l'utilisation de cette option et de la méthode des moindres carrés pourrait être admise et utilisée comme une boîte noire au même titre que la fonction *alea* ou *random* donne des nombres au hasard).

Plus généralement, les élèves doivent avoir acquis l'idée que, lors d'une simulation, lorsqu'on multiplie les échantillons, la fréquence moyenne se stabilise vers la probabilité. Ce qui pourrait situer cette expérience au niveau de la classe de Première.

5.3. Étape n^{o} 1 : utilisation de textes historiques et d'un tableur pour déterminer des relations entre des mesures observées

Mon exemple est choisi dans l'histoire de l'astronomie, science très liée aux mathématiques et dans laquelle on utilise beaucoup les statistiques et les probabilités pour conjecturer de grandes lois. En particulier, un problème essentiel et constant est de déterminer si, dans l'étude d'un modèle, un petit écart entre la moyenne des mesures observées et la valeur théorique est dû à des erreurs d'observations ou à des causes naturelles non encore prises en compte dans le modèle. Les travaux cités de Laplace avaient cette finalité.

5.3.1. Découvrir une relation fonctionnelle à partir de relevés de coordonnées de points. Exemple de la troisième loi de Kepler

Voici un exemple important de la démarche scientifique : l'astronome Tycho Brahé (1546-1601) fait de nombreuses observations très précises sur les trajectoires des planètes ; Kepler (1571-1630) utilise les résultats de ces observations pour vérifier la théorie de Copernic suivant laquelle les planètes tournent autour du Soleil d'un mouvement circulaire, uniforme et centré au Soleil. En particulier, une petite différence de 8 minutes d'angles entre les observations de Mars par Tycho Brahé et la position calculée de cette planète sur le cercle théorique conduit Kepler à rejeter le modèle circulaire de Copernic et à trouver ses deux premières lois : « Les orbes des planètes sont des ellipses dont le centre du soleil occupe l'un des foyers. Les aires décrites autour de ce centre, par les rayons vecteurs des planètes sont proportionnelles aux temps employés à les décrire » (Laplace, 1836, p. 113).

Ces lois seront démontrées par Newton (entre 1666 et 1687) à partir du principe de l'attraction universelle, après l'invention de la dynamique et du calcul infinitésimal.

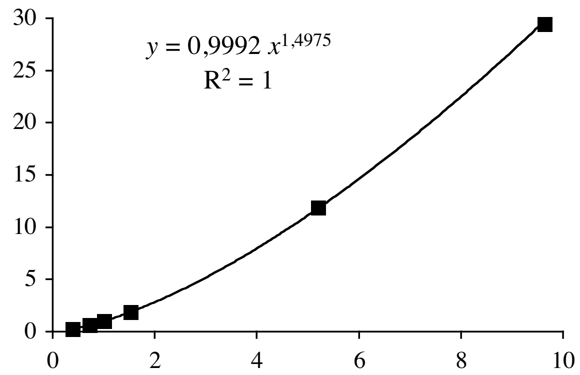
Comment faire découvrir la troisième loi de Kepler par les élèves ? Laplace nous rappelle aussi que, « après un grand nombre de tentatives continuées pendant dix-sept ans, il [Kepler, *L'Harmonie du monde*, 1619] reconnut enfin, que les carrés des temps de révolution des planètes, sont comme les cubes des grands axes de leurs orbites » (Laplace, 1836, p. 114).

C'est la troisième loi de Kepler. Elle fut ensuite vérifiée et confirmée lors de la découverte des satellites de Jupiter par Galilée. Voici un tableau sur lequel Kepler aurait pu travailler :

Planètes	Distance moyenne au Soleil : a	Durée moyenne des révolutions : T	T^2 / a^3
Mercure	0,39	0,24	0,978
Vénus	0,72	0,62	1,014
la Terre	1,00	1,00	1,000
Mars	1,52	1,88	1,008
Jupiter	5,20	11,86	1,001
Saturne	9,64	29,46	0,969

Suggestions d'activités :

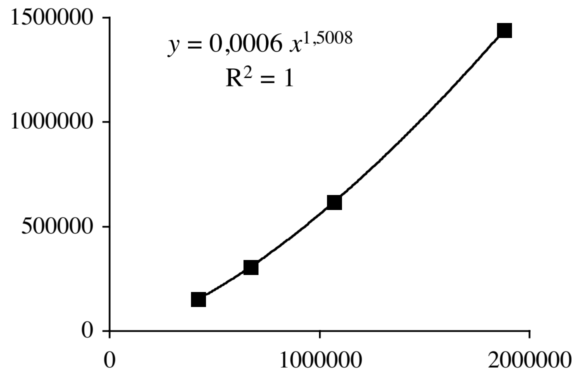
- À partir des données des trois premières colonnes du tableau ci-dessus, faire remplir par les élèves les colonnes des valeurs T^i/a^j pour i et $j = 1, 2$ ou 3 et leur demander de proposer un choix (une formule algébrique, une fonction) justifié pour le modèle. On pourra représenter T en fonction de a et utiliser le tableur pour obtenir des courbes de tendances. On obtiendra, par exemple, $y = 0{,}9992\ x^{1{,}4975}$ et le graphique suivant :



- Vérifier si le modèle fonctionnel choisi ci-dessus est adapté aux satellites de Jupiter, découverts par Galilée en 1610. Vous utiliserez les données suivantes recueillies pour les satellites de Jupiter. (Vous pouvez, par exemple, entrer a et calculer T puis comparer avec le tableau.)

Satellite de Jupiter	Distance du satellite à Jupiter (en km) : a	Période de révolution du satellite autour de Jupiter (en s) : T
Io	$4{,}22.10^5$	$1{,}53.10^5$
Europe	$6{,}71.10^5$	$3{,}07.10^5$
Ganymède	$1{,}07.10^6$	$6{,}19.10^5$
Callisto	$1{,}88.10^6$	$1{,}44.10^6$

- Vous pouvez alors changer la fonction trouvée pour les planètes et obtenir la nouvelle expression pour les satellites de Jupiter. Ce qui donne $T^2/a^3 = 3{,}11491.10^{-7}$ et le graphique :



5.3.2. Étude, par les élèves, d'un extrait du texte de Cournot

Il s'agit de l'*Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*[4] (1851). En particulier, on posera les questions suivantes aux élèves :

- À partir de l'étude du paragraphe 43 du texte de Cournot, énoncer la loi de Bode, et construire une feuille de calcul dans un tableur qui permette la vérification, ou non, de cette loi.
- Faire une synthèse des paragraphes 44 à 46 du texte de Cournot.

5.4. Étape n° 2 : utilisation de la simulation pour conjecturer les premières règles du calcul des probabilités

En utilisant les annexes 1 et 2, nous pouvons comparer les définitions de Bayes (1763) et l'axiomatique de Kolmogorov (1933). Pour ces deux théories, une même formule permet de calculer la probabilité de l'événement « A ou B » lorsque les deux événements sont incompatibles. Or, en dehors d'exemples triviaux, la théorie du chaos nous laisse supposer que cette notion d'incompatibilité relève quelque peu de l'incantation magique. Du point de vue didactique, il nous semble donc plus naturel de commencer par une introduction expérimentale de la formule $p(\text{A ou B}) = p(\text{A}) + p(\text{B}) - p(\text{A et B})$ au moyen d'une simulation.

En classe, on peut commencer par faire étudier le texte de Bayes cité en annexe 1, puis on utilise l'ordinateur pour simuler l'état de marche ou de panne d'un système formé de deux pièces montées en parallèle : le système marche si l'une ou l'autre des pièces est en état de marche[5].

5.4.1. Premier résultat simulé et observé

Sur la feuille de calcul proposée, on peut faire varier les valeurs de p(A) et celles de p(B). Dans chacune des dix cellules Mi, on affiche la moyenne de 1000 tirages simulés. Voici un exemple de résultats ainsi obtenus :

p(A) =	0,1		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
p(B) =	0,5		0,56	0,55	0,56	0,55	0,54	0,53	0,57	0,54	0,53	0,54

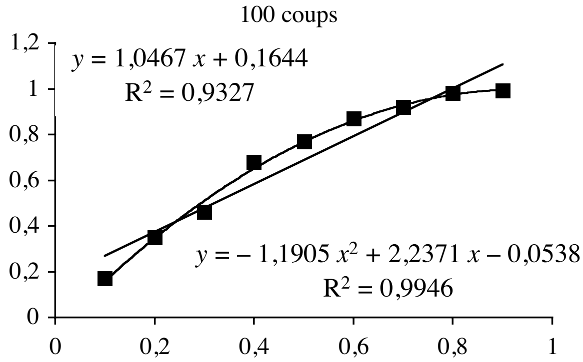
Moyenne théorique = 0,55 Moyenne observée = 0,55

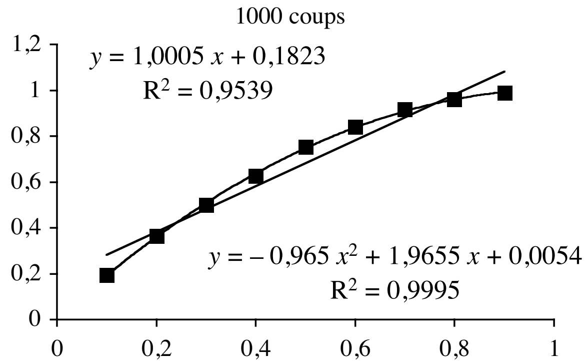
L'objectif est de faire découvrir une loi, une formule, qui permette de calculer p(A ou B) connaissant p(A) et p(B). Les élèves doivent commencer à se poser les bonnes questions s'ils comparent les résultats de la simulation à la formule démontrée par Bayes. La valeur théorique est donnée à titre indicatif ; on peut la supprimer pour un usage en classe.

5.4.2. Second résultat simulé et observé

Sauf si un élève connaît déjà le résultat, il est très peu probable qu'il conjecture directement que p(A ou B) = p(A) + p(B) – p(A et B), d'où l'activité suivante où l'on va se placer dans une situation plus simple en choisissant l'égalité des probabilités de A et de B. Dans ce cas, nous essayons de déterminer une loi en utilisant les équations des courbes de tendances du 1er et du 2^{e} degrés. Pour cela, on simule 1000 essais pour chaque valeur théorique de p(A) = 0,1 ; 0,2 ;... ; 0,9. On construit deux graphiques (pour 100 essais, puis pour 1000 essais) avec les équations des courbes d'ajustements. Il faut faire découvrir que $p = 2p(\text{A}) - p(\text{A})^2$, ou encore que $f(x) = 2x - x^2$. Voici un exemple de résultats obtenus de cette manière :

p(A) = x	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
100 coups	0,17	0,35	0,46	0,68	0,77	0,87	0,92	0,98	0,99
1000 coups	0,194	0,366	0,501	0,626	0,751	0,841	0,916	0,960	0,988



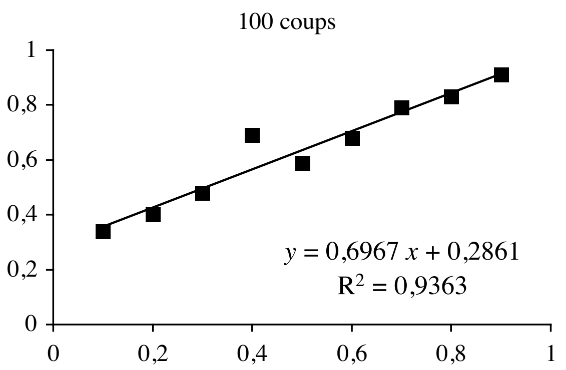


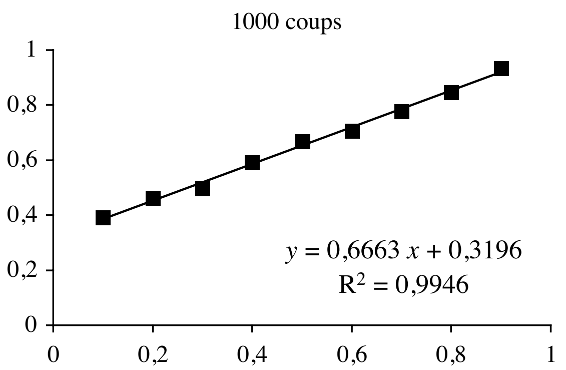
5.4.3. Troisième résultat simulé et observé

On essaie maintenant de déterminer la loi lorsque p(A) est fixé et p(B) variable. On utilise les équations des courbes de tendance du 1er degré. Pour cela, on simule 1000 essais pour chaque valeur de p(B) variant de 0,1 à 0,9, par pas de 0,1. On peut faire varier la valeur de p(A). Par exemple, pour p(A) = 0,3, on obtient :

p(B) = x	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
100 coups	0,34	0,40	0,48	0,69	0,59	0,68	0,79	0,83	0,91
1000 coups	0,391	0,463	0,496	0,592	0,669	0,706	0,778	0,847	0,933

On construit deux graphiques (pour 100 essais, puis pour 1000 essais) avec les équations des courbes d'ajustements. Maintenant, il s'agit de parvenir à conjecturer que $p(\text{A ou B}) = p(\text{A}) + p(\text{B}) - p(\text{A})\,p(\text{B})$, ou encore que $f(x) = p(\text{A}) + x - p(\text{A})x = (1 - p(\text{A}))x + p(\text{A})$.





Conclusions provisoires

Si l'usage de l'ordinateur dans une classe se résume, pour l'élève, à regarder une simulation entièrement préparée par le professeur, nous ne faisons qu'une partie de notre travail d'enseignant. On risque de tromper les élèves en leur faisant croire qu'ils vont pouvoir utiliser les TIC et les maîtriser si, en même temps, on ne leur donne pas le temps et le matériel nécessaires à tout apprentissage. Non, l'utilisation raisonnée de l'outil informatique n'est pas innée.

Faire de la simulation, ce n'est pas aussi facile que de manipuler une console de jeu vidéo. Tant qu'il n'y aura pas d'heures explicitement prévues dans l'emploi du temps de l'élève pour qu'il puisse travailler sur ordinateur, il n'y aura pas formation de sens.

Mais cet article est écrit pour faire de la prospective. Donnons-nous les moyens, en temps et en matériels, pour conduire au lycée une véritable initiation à la recherche scientifique. Si les horaires baissent, il faudra assumer : le risque sera grand de voir le professeur revenir à un enseignement *ex cathedra* pour « boucler » le programme. Ce serait en contradiction avec l'esprit des nouveaux contenus et des nouveaux objectifs. Et ce serait regrettable car, dans ce nouveau cadre, l'utilisation de l'outil informatique, associée à l'étude de textes historiques, aurait pu permettre de réconcilier bon nombre de nos élèves avec les sciences en général et les mathématiques en particulier.

Annexe 1

Thomas Bayes, *Essai en vue de résoudre un problème de la doctrine des chances* (rééd. 1988, extraits)

Section 1.
Définitions :
Plusieurs événements sont incompatibles lorsqu'ils sont tels que si l'un se produit, aucun autre ne le peut.
Deux événements sont contraires, lorsque l'un ou l'autre doit arriver sans qu'ils ne le puissent tous les deux à la fois.
La probabilité d'un événement quelconque est le rapport entre la valeur à laquelle on devrait estimer une espérance dépendant de la réalisation de cet événement et la valeur de la chose espérée s'il se réalise.
J'entends par la « chance » la même chose que la « probabilité ».
Des événements sont indépendants quand le fait que l'un se produise n'accroît ni ne diminue la probabilité des autres.

Proposition 1.
Quand plusieurs événements sont incompatibles, la probabilité pour que l'un ou l'autre d'entre eux se réalise est la somme des probabilités de chacun d'entre eux.
Démonstration : supposez qu'il y ait trois événements de ce type ; que, quel que soit celui qui se produise, je doive recevoir N et que les probabilités du 1er, du 2d et du 3ème soient respectivement a/N, b/N, c/N. Alors par la définition des probabilités la valeur de mon espérance par le 1er sera a, par le 2d sera b, et par le 3ème sera c. De là, la valeur de mes espérances pour tous les trois sera $a + b + c$. Mais la somme de mes espérances à partir de tous les trois est dans ce cas une espérance de recevoir N

dans le cas où l'un ou l'autre d'entre eux se produit. De là, (par la définition 5), la probabilité de l'un ou l'autre d'entre eux est : $(a + b + c)/N$ ou $a/N + b/N + c/N$. Soit la somme des probabilités de chacun d'entre eux.

[...] Proposition 3.
La probabilité que deux événements subséquents se produiront tous les deux est en raison composée de la probabilité du premier et de la probabilité du second en supposant que le premier se produise.
[Soit, en langage moderne : $p(\text{A et B}) = p(\text{A}).p(\text{B/A})$.]

[...] Proposition 6.
La probabilité pour que plusieurs événements indépendants arrivent tous est en raison composée des probabilités de chacun.
[Soit, en langage moderne : si deux événements sont indépendants, alors $p(\text{A et B}) = p(\text{A}).p(\text{B})$.]

Annexe 2

La théorie axiomatique de Kolmogorov (1933)

Dans un manuel scolaire destiné aux élèves de Terminales CE des lycées (Gautier *et al*., 1975, p. 355), on en trouve cette présentation : « Étant donné un espace probabilisable fini (E, B(E)), on appelle probabilité sur (E, B(E)) toute application p de B(E) dans l'ensemble des réels satisfaisant aux axiomes suivants :
1. $p(\text{E}) = 1$
2. Pour tout élément A de B(E), $p(\text{A}) \geq 0$
3. Si A et B sont deux événements incompatibles alors $p(\text{A ou B}) = p(\text{A}) + p(\text{B})$.
Le triplet (E, B(E), p) est alors un espace probabilisé fini. »

À l'université elle entre dans le cadre de la théorie de la mesure (Henry, 1999, p. 18) : « Une probabilité est une mesure sur un ensemble Ω, application σ-additive d'une tribu de parties de Ω dans $[0 ; 1]$. »

De ces axiomes, on déduit les règles de base du calcul des probabilités et, en particulier : $p(\text{A ou B}) = p(\text{A}) + p(\text{B}) - p(\text{A et B})$. Puis vient le théorème des probabilités composées : $p(\text{A et B}) = p(\text{A}).p(\text{B/A})$. Enfin, on définit l'indépendance de deux événements par la relation $p(\text{A et B}) = p(\text{A}).p(\text{B})$.

On remarquera une différence entre les deux introductions au calcul des probabilités : la définition de l'indépendance de deux événements est une définition a priori chez Bayes, alors qu'elle est a posteriori chez Kolmogorov. Cette notion d'indépendance est un problème difficile de didactique.

Références bibliographiques

BAYES Thomas (1988), *Essai en vue de résoudre un problème de la doctrine des chances*, trad. de J.-P. Cléro in *Cahiers d'histoire et de philosophie des sciences*, n° 18 (1re éd. : 1763).

BERNOULLI Jacques (1987), *Ars Conjectandi*, trad. par N. Meusnier, suivi de la *Lettre à un ami sur les parties du jeu de paume*, Rouen, IREM (1re éd. : 1713).

BORDIER Jacques (1991), *Un modèle didactique utilisant la simulation sur ordinateur pour l'enseignement de la probabilité*, thèse de doctorat, université Paris 7-Denis Diderot.

COURNOT Antoine Augustin (1851), *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*, Paris, Hachette.

GAUTIER C., GIRARD G., GERLL D., THIERCÉ C. et WARUSFEL A. (1975), *Aleph1. Analyse. Terminales CE*, Paris, Hachette.

HENRY Michel (1999), « L'introduction des probabilités au lycée », *Repères-Irem*, n° 36.

LAPLACE Pierre-Simon, (1812), *Théorie analytique des probabilités*, Paris, Courcier.

LAPLACE Pierre-Simon, (1836), *Exposition du système du monde*, Paris, Bachelier (1re éd. : 1796).

KEPLER Jean (1979), *L'Harmonie du monde*, Paris, Blanchard (1re éd. : 1619).

NAMER Émile (1979), *Le Beau Roman de la physique cartésienne et de la science exacte de Galilée*, Paris, Vrin.

Notes

1. Cet article est issu des travaux en cours du groupe de recherche-action « Modélisation et simulation » de l'IREM de la Réunion, regroupant Jean-Claude Lise, Michel Gontier et moi-même.
2. Des extraits de ce texte peuvent être téléchargés sur le site de l'IREM de la Réunion : http://www.reunion.iufm.fr/recherche/irem/divers/.
3. Des extraits de ce texte se trouvent sur le site de l'IREM (voir note 2).
4. Des extraits de ce texte se trouvent sur le site de l'IREM (voir note 2).
5. Le fichier Excel correspondant (probastat.xls) est à télécharger sur le site de l'IREM (voir note 2).