INTRODUCTION AUX LOIS DE PROBABILITES CONTINUES : PROBLEMES EPISTEMOLOGIQUES

Jean Claude THIENARD
Irem de Poitiers

Cet article fait suite à une réflexion menée par l'auteur lors de l'une des rencontres de l'Atelier de Culture Scientifique de l'Irem de Poitiers. Ce groupe est composé majoritairement d'enseignants de mathématiques mais aussi de physiciens et de philosophes. Son objectif essentiel est de faire en sorte que les enseignants d'une discipline scientifique comprennent l'épistémologie, les enjeux et les méthodes des autres disciplines scientifiques afin de tisser des liens pertinents entre ces différentes disciplines.

La notion de lois de probabilités continues à densité a été introduite dans le programme de terminale S :

- *Calcul de probabilités d'intervalles pour des lois de probabilités à densité.*
- *Exemples de lois continues : lois continues à densité — loi uniforme sur* [0, 1] *— loi de durée de vie sans vieillissement.*

  Application à la désintégration radioactive. Loi exponentielle de désintégration des noyaux.

Cette introduction ne va pas de soi, à moins de décider de s'en tenir à un aspect purement formel juste suffisant pour faire ronronner les élèves sur les quelques situations scolaires susceptibles d'enrichir les annales du BAC [1].

Le passage de modélisations dans le monde des probabilités finies aux modélisations dans le monde des probabilités continues pose de redoutables problèmes qui, s'ils ne sont pas repérés, risquent de générer autant d'obstacles didactiques et d'empêcher de façon durable la constitution d'un savoir vrai sur le sujet c'est-à-dire d'un véritable outil de pensée [2].

1 La nécessité de sortir de la vision des mathématiques offerte aux élèves par le cadre scolaire classique est chaque jour plus évidente, elle est attestée par la remise en cause de la légitimité de cet enseignement. Cette discipline hautement intellectuelle, formatrice de l'esprit, pensée structurante à l'œuvre dans les activités humaines les plus variées, que ce soit au niveau de maintes recherches fondamentales ou d'ingénieries quotidiennes doit au niveau de l'enseignement être montré pour ce qu'elle est : un outil de pensée ; cela oblige à ne pas cantonner les élèves dans le cadre étriqué des exercices scolaires destinés à faire acquérir la maîtrise de quelques techniques.

2 Cela doit constituer l'objectif permanent d'un enseignement des mathématiques.

L'article n'a pas pour visée de critiquer le programme et ses commentaires [3] mais de mettre en évidence les difficultés rencontrées lors d'une première approche de la notion de probabilité continue, en analysant celles rencontrées par les pionniers : rencontres de situations conduisant à des paradoxes, résultats en désaccord manifeste avec l'expérience… et de mettre en évidence qu'elles ont pour source ce que depuis Bachelard on appelle des *obstacles épistémologiques.*

Les obstacles épistémologiques sont aujourd'hui occultés par la présentation axiomatique de la théorie et se sont mués en obstacles didactiques car, si la notion de loi continue est aujourd'hui parfaitement définie, son articulation au monde des probabilités finies et au monde de l'expérience reste problématique pour qui y est confronté pour la première fois. La compréhension des notions qui font la matière d'un cours renvoie à un ensemble de questions — Pourquoi ces définitions ? Comment modéliser dans le continu ?… — qu'il ne faut pas éluder si l'on souhaite que l'élève acquière un savoir utile c'est-à-dire un savoir qui sert à penser.

Le fil conducteur de l'étude est constitué de quelques textes extraits du *Calcul des Probabilités* (1889) de J. Bertrand professeur au Collège de France. Ces textes sont, pour certains d'entre eux, bien connus et sont souvent cités sous le titre *Les paradoxes de Bertrand.*

La première série de textes se rapporte aux paradoxes de Bertrand ; ceux-ci sont en général bien connus dans la mesure où la littérature sur le calcul des probabilités et sur son enseignement y fait souvent référence. Ils sont alors montrés ou comme des curiosités ou pour attester de la difficulté historique à définir des probabilités dans le continu. La thèse développée dans l'article est que ces paradoxes montrent beaucoup plus que cela et révèlent ce que Bachelard a appelé un obstacle épistémologique. A ce titre ils peuvent donner à réfléchir à tous ceux qui sont confrontés à la tâche de dispenser un premier enseignement sur le sujet des probabilités continues.

La deuxième série de textes a pour objet la critique par Bertrand de modélisations faites dans le continu, notamment de celle faite par Maxwell dans le cadre de la théorie cinétique des gaz. Ces textes soulèvent un autre type de difficulté lié à l'introduction des probabilités dans le continu. Deux de ces textes seront examinés afin :

- de dégager les rapports entre observations et hypothèses de modélisation
- de poser le problème de l'adéquation entre prédictions et résultats expérimentaux
- de mettre en évidence les lacunes qui peuvent conduire à l'impossibilité de comprendre le hiatus entre ce que dit la théorie et ce que dit le bon sens le plus élémentaire.

I. — Le cadre théorique du traité de Bertrand

Le traité s'ouvre sur la définition de la probabilité : *« La probabilité est estimée par l'énumération des cas favorables, rapprochée à celle des cas possibles. »* [4]

3 Très critiquables au demeurant ; notamment par la volonté affichée par les commentaires d'occulter la notion de variable aléatoire en introduisant celle de probabilité d'un intervalle. [0, t[n'est pas un événement et p([0, t[)n' a pas de sens alors que si X est une durée de vie par exemple $0 \le X \le t$ est un événement et que l'écriture $p(0 \le X \le t)$ a un sens immédiat qui est en relation avec tout ce que l'élève connaît déjà du calcul des probabilités.

4 Cette définition est empruntée de P.S. Laplace : Introduction de la Théorie Analytique des Probabilités [2] et [3]. Suite à cette définition Laplace explique que la définition ne vaut que si les divers cas sont également possibles – comme dans l'urne, outil universel de modélisation – et que sinon on commence par calculer les probabilités des divers cas !

Cette définition vaut dans le cas des probabilités finies. Elle est évidemment circulaire [5] puisqu'elle suppose que tous les cas doivent être également possibles et suppose donc une définition de l'équiprobabilité ! Elle renvoie de façon implicite [6] — pour celui qui sait — au paradigme de l'urne, c'est-à-dire à la méthode élaborée par les créateurs du calcul pour rendre compte — pour modéliser dit-on aujourd'hui — d'un certain type d'expériences dont les résultats dépendent *du hasard.*

II. — Texte I : Choisir un nombre au hasard. Extrait du *Calcul des Probabilité* (1889) de J. Bertrand [4]

4. Une remarque encore est nécessaire : l'infini n'est pas un nombre ; on ne doit pas, sans explication, l'introduire dans les raisonnements. La précision illusoire des mots pourrait faire naître des contradictions. Choisir au hasard, entre un nombre infini de cas possibles, n'est pas une indication suffisante.

On demande, par exemple, la probabilité pour qu'un nombre, entier ou fractionnaire, commensurable ou incommensurable, choisi au hasard entre 0 et 100 soit plus grand que 50. La réponse semble évidente : le nombre des cas favorables est la moitié de celui des cas possibles. La probabilité est 1/2.

Au lieu du nombre, cependant, on peut choisir son carré. Si le nombre est compris entre 50 et 100, le carré le sera entre 2500 et 10000.

La probabilité pour qu'un nombre choisi au hasard entre 0 et 10000 surpasse 2500 semble évidente : le nombre des cas favorables est les trois quarts du nombre des cas possibles. La probabilité est 3/4.

*Les deux problèmes sont identiques. D'où vient la différence des réponses? **Les énoncés manquent de précision** [7].*

Les contradictions de ce genre peuvent être multipliées à l'infini.

Commentaires

Le problème posé n'est pas du ressort du calcul des probabilités ; choisir un nombre réel au hasard ne correspond à aucune expérience réalisable [8]. Il ne s'agit là au plus que d'une expérience de pensée qui conduit à traiter un problème qui relève de la théorie de la mesure sur le mode métaphorique des probabilités [9]. Admettons que le problème relève du calcul des probabilités. Le *référent* de Bertrand pour avancer sa solution est le *modèle de l'urne* qui s'impose comme le seul possible par analogie avec le cas des probabilités finies [10].

- « *La réponse semble **évidente** : le nombre des cas favorables est la moitié de celui des cas possibles. La probabilité est 1/2.* »
- *« La probabilité pour qu'un nombre choisi au hasard entre 0 et 10000 surpasse 2500*

5 Pour une étude plus complète, se reporter à : Enseigner les probabilités au lycée, Article A propos de la définition de la probabilité. Publication de la commission inter IREM statistiques et probabilités. [6]

6 Implicite, là gît le problème, car cette définition n'a de sens qu'en relation avec une modélisation par l'urne. Détachée de ce contexte fondateur de toute démarche probabiliste, la définition n'est pas opératoire ce qui apparaît dans la suite du traité de Bertrand où pour instituer le calcul il est obligé de développer toute une série d'exemples qui seront autant de modèles à imiter lors de la résolution d'exercices.

7 Tout ce qui est surligné en gras dans les textes originaux est le fait de l'auteur de l'article.

8 Une machine, une table de nombre au hasard peut permettre de générer des nombres au hasard à un nombre de décimales fixé à l'avance, mais jamais un nombre réel.

9 Choisir au hasard un nombre entre 0 et 100 s'écrivant avec n décimales relève du calcul des probabilités. On peut , en effet composer une urne qui contient tous ces nombres et effectuer un tirage dans cette urne !

10 Ce mode analogique est déjà celui utilisé par Buffon dans la solution qu'il apporte aux problèmes du jeu du Franc Carreau et de l'aiguille de Buffon. On peut se référer à l'article de Michel Henry paru dans Repère Irem N°53, 2003. [9]

*semble **évidente** : le nombre des cas favorables est les trois quarts du nombre des cas possibles. La probabilité est 3/4.* »

Les nombres sont implicitement pensés comme éléments d'une urne contenant une infinité d'éléments. Il y a ainsi l'urne des nombres et celle de leurs carrés et un tirage dans ces urnes conduit, selon la règle du paradigme de l'urne (nombre de cas favorables sur nombre de cas possibles), aux résultats énoncés. Cette démarche a une implication très forte, qui traduite en termes modernes, s'exprime par le fait que toute variable aléatoire définie dans le continu suit une loi uniforme, ce qui ne peut conduire qu'à des paradoxes lorsque les variables considérées ne sont pas indépendantes.

En effet si X est la variable aléatoire « valeur du nombre tiré » (choisi au hasard) suivant une loi uniforme, on a alors $P(X < x) = x/100$ pour $x \in [0, 100]$, 0 si $x < 0$ et 1 si $x > 100$, ce qui est une manière métaphorique raisonnable de traduire l'équirépartition des réels dans [0, 100], $P(X < x)$ étant alors la mesure usuelle sur $\mathbf{R}$.

On a alors :

$$P(X^2 < x) = P(X < \sqrt{x}) = \sqrt{x}/100 ,$$

ce qui montre que la loi suivie par X^2 n'est pas uniforme. Il vient alors :

$$P(X^2 \geq 2500) = 1 - P(X^2 < 2500) =$$

$$1 - \sqrt{2500}/100 = 1/2 = P(X > 50) ,$$

le paradoxe a disparu.

La conclusion de Bertrand « *Les deux problèmes sont identiques. D'où vient la différence des réponses? Les énoncés manquent de précision.* » est inexacte. Ce n'est pas un manque de précision *sur choisir au hasard entre un nombre infini de cas possibles qui* génère un paradoxe. Le paradoxe est généré par l'usage implicite qui est fait du paradigme de l'urne : choisir au hasard c'est effectuer un tirage dans une urne, or ici :

- l'urne est, comme on l'a vu, humainement impossible à composer
- dans l'hypothèse où l'on se prend pour Dieu acceptant de jouer avec cette chose très humaine que sont les réels, que l'on voit tous les nombres entre 0 et 100 parfaitement équirépartis disposés devant soi sur une droite numérique et que l'on puisse les yeux bandés en choisir un au hasard, l'équirépartition des nombres réels conduit à affirmer que $P(X < x) = x/100$ pour $x \in [0, 100]$, 0 si $x < 0$ et 1 si $x > 100$, et donc que $P(X > 50) = 1/2$; Dieu verra alors les nombres étiquetés comme carrés, par un étiquetage qui restera probablement à jamais à inventer pour les hommes, disposés sur la même droite numérique mais non équirépartis [11], autrement dit que la répartition des carrés n'est pas uniforme [12].
- Lorsque l'on passe du fini à l'infini un autre type d'obstacle surgit [13] : le passage de la *cardinalité* à la *mesure*. La référence au modèle de l'urne crée une référence à la cardinalité, or les premières probabilités calculées dans le continu (voir exemples du jeu du franc carreau et de l'aiguille de Buffon) s'inscrivent dans le cadre de la mesure : elles sont calculées par des rapports d'aires. Cette mesure ne peut être pensée qu'uniforme : *« le nombre d'éléments »* est proportionnel à la surface qui les contient, selon le principe implicite que si une boite contient n boules (identiques) alors une boîte

11 Ceci signifie que $mes\{t,\ t^2 < x\} \neq x/100$, $t \in [0; 100]$; plus simplement les carrés de [0; 0,1] sont répartis dans [0; 0,01] et ceux de [0,1; 0,2] sont répartis dans [0,01; 0,04],… ce qui suffit à attester de ce qui précède.

12 Pour les hommes, le lien entre X et X^2 rend compte de cela.

13 Lié au précédent.

de volume double contiendra 2*n* boules. Ici encore la référence implicite à l'urne conduit à la référence implicite à la cardinalité et donc à la mesure uniforme. *Cette référence se constitue en obstacle épistémologique prégnant ainsi que l'atteste l'exemple suivant.*

III. — Texte II : Choisir une corde au hasard. Extrait du *Calcul des Probabilités* (1889) de J. Bertrand [4]

5. On trace au hasard une corde dans un cercle. Quelle est la probabilité pour qu'elle soit plus petite que le côté du triangle équilatéral inscrit?

On peut dire : si l'une des extrémités de la corde est connue, ce renseignement ne change pas la probabilité ; la symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

L'une des extrémités de la corde étant connue, la direction doit être réglée par le hasard. Si l'on trace les deux côtés du triangle équilatéral ayant pour sommet le point donné, ils forment entre eux et avec la tangente trois angles de 60°. La corde, pour être plus grande que le côté du triangle équilatéral, doit se trouver dans celui des trois angles qui est compris entre les deux autres. La probabilité pour que le hasard entre trois angles égaux qui peuvent le recevoir le dirige dans celui-là semble, par définition, égale à 1/3.

On peut dire aussi : si l'on connaît la direction de la corde, ce renseignement ne change pas la probabilité. La symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

La direction de la corde étant donnée, elle doit, pour être plus grande que le côté du triangle équilatéral, couper l'un ou l'autre des rayons qui composent le diamètre perpendiculaire, dans la moitié la plus voisine du centre. La probabilité pour qu'il en soit ainsi semble, par définition, égale à 1/2.

On peut dire encore : choisir une corde au hasard, c'est en choisir au hasard le point milieu. Pour que la corde soit plus grande que le côté du triangle équilatéral, il faut et il suffit que le point milieu soit à une distance du centre plus petite que la moitié du rayon, c'est-à-dire à l'intérieur d'un cercle quatre fois plus petit en surface. Le nombre des points situés dans l'intérieur d'une surface quatre fois moindre est quatre fois moindre. La probabilité pour que la corde dont le milieu est choisi au hasard soit plus grande que le côté du triangle équilatéral semble, par définition, égale à 1/4.

Entre ces trois réponses, quelle est la véritable? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée.

Commentaires

- Ce problème peut relever du calcul des probabilités, on peut en effet imaginer des expériences qui conduisent à déterminer « au hasard » une corde sur un cercle [14] selon les procédures imaginées par Bertrand.
- Dans la première procédure la corde est définie par l'angle θ qu'elle fait avec la tangente, cet angle est une variable aléatoire [15]. Pour Bertrand, il est **évident** — le modèle de l'urne agissant de la même façon que précé-

14 On matérialise un cercle sur le sol et l'on fait rouler en aveugle un bâton cylindrique dont la position d'arrêt donnera la droite qui portera la corder éventuelle ou l'on jette en aveugle une bille dont le point d'arrêt donnera le centre d'une corde éventuelle….

15 La notion de variable aléatoire n'a pas encore été pensée à l'époque de Bertrand. Il faudra attendre les année 1920 pour qu'elle émerge suite aux travaux de P. Lévy et de Khintchine.

demment — qu'elle suit, ce qu'aujourd'hui on appelle, une loi uniforme sur $[0; \pi]$, (en tenant compte de la symétrie du cercle on la considérera définie sur $[0 ; \pi/2]$ afin de simplifier les calculs).

Cela signifie que $P(\theta < x) = 2x/\pi$, pour $x \in [0, \pi/2]$, 0 si $x < 0$ et 1 si $x > \pi/2$; le résultat énoncé par Bertrand est alors :

$$P(\theta > \pi/3) = 1 - P(\theta < \pi/3) = 1 - \frac{\pi / 3}{\pi / 2} = 1/3 .$$

• Dans la deuxième procédure la corde est définie par sa distance d au centre qui est une variable aléatoire. Pour Bertrand, il est **évident** — le modèle de l'urne agissant de la même façon que précédemment — qu'elle suit une loi uniforme sur $[0 ; 1]$; cela signifie que $P(d < x) = x$ pour $x \in [0, 1]$, 0 si $x < 0$ et 1 si $x > 1$; le résultat énoncé par Bertrand est alors $P(d < 1/2) = 1/2$.

Le paradoxe provient de ce que si l'on suppose que la loi de θ est uniforme alors celle de d ne peut pas être uniforme, en effet : $d = cos\theta$,

$$P(d < x) = P(cos\theta < x) =$$

$$P(\theta > \mathrm{Arc}cos\ x) = 1 - \frac{2Arc\cos x}{\pi},$$

d suit donc une loi de densité $\dfrac{2}{\pi\sqrt{1 - x^2}}$.

On retrouve alors le résultat énoncé par Bertrand sous l'hypothèse que θ est une variable qui suit une loi uniforme sur $[0, \pi/2]$:

$$P(d < 1/2) = 1 - \frac{2Arc\cos 1/2}{\pi} = 1/3 .$$

Si l'on fait l'hypothèse que d suit une loi uniforme sur $[0 ; 1]$, $P(d < x) = x$, $d = cos\theta$,

$$P(\theta < x) = P(\mathrm{Arc}cos\ d < x) =$$

$$P(d > cosx) = 1 - P(d < cosx) = 1 - cosx$$

pour $x \in [0, 1]$, θ suit une loi de densité $sin\ x$.

On retrouve alors le résultat énoncé par Bertrand, sous l'hypothèse que d est une variable qui suit une loi uniforme sur $[0 ; 1]$:

$$P(\theta > \pi/3) = 1 - P(\theta < \pi/3) = 1 - (1 - cos\pi/3) = 1/2 .$$

• La conclusion tirée par Bertrand : « *Entre ces trois réponses, quelle est la véritable? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée.* » est ambiguë. Que signifie la question est mal posée ? En fait la question n'est pas posée puisque la question n'a de sens que par rapport à une expérience aléatoire qui détermine la corde ; or la corde est choisie au hasard indépendamment de toute expérience aléatoire, cela laisse le problème totalement indéterminé et donc sans solution ! Bertrand semble cependant indiquer par *la question est mal posée* qu'il suffirait de connaître le bon choix du paramètre aléatoire adéquat θ, d… pour pouvoir décider de la bonne solution. Ceci apparaît *évident*… si l'on raisonne dans le cadre du paradigme de l'urne.

• En fait la solution du problème étant liée — il en est ainsi de tout problème de probabilité — à une expérience aléatoire, peu importe alors — même s'il y en a une qui est plus commode que les autres — que ce soit la variable θ ou d qui soit choisie pour traiter le problème puisqu'elles sont fonctionnellement dépendantes[16]. Ce qui importe est de connaître la loi de répartition ou la densité de probabi-

16 Si θ suit une loi de densité f sur $[0; \pi/2]$, $P(\theta < x) = \int_0^x f(t)dt$ et comme $d = \cos\theta$:

$P(d < x) = P(\cos\theta < x) = P(\theta > \mathrm{Arccos}\ x) =$

$$1 - P(\theta < \mathrm{Arccos}\ x) = 1 - \int_0^{Arc\cos x} f(t)dt = \int_0^x \frac{f(Arc\cos u)}{\sqrt{1 - u^2}}du$$

qui donne la relation entre la densité de θ et celle de d.

lité de l'une d'entre elles ; rien ne peut être alors posé a priori, seule l'expérience répétée peut conduire au choix d'un modèle et il y a peu de chance de tomber sur une loi uniforme pour d ou pour θ !

Quelques conclusions

Les premières modélisations d'expériences aléatoires ayant une infinité de possibles — jeu du franc carreau, aiguille de Buffon [17] — sont faites sur le modèle de l'urne. Ceci implique que les paramètres aléatoires choisis pour décrire les évènements auxquels on cherche à assigner une probabilité ne peuvent que suivre des lois de distribution uniforme [18].

Les paradoxes de Bertrand attestent de la prégnance de ce modèle et de l'obstacle épistémologique qu'il crée.

Ils attestent également d'une perte de sens qui s'est opérée — probablement à partir du grand Traité de Laplace *Théorie analytique des probabilités (1812)* [3] — depuis que l'on a donné des exposés purement mathématique du calcul des probabilités afin d'en asseoir les bases et d'en expliciter les règles. Ces exposés donnent la syntaxe du calcul — qui résulte de calculs sur des modèles d'urne- et en occultent la sémantique : l'urne n'est jamais donnée mais toujours à construire [19]. En effet le calcul des probabilités s'applique, les situations dont il traite appartiennent au monde « réel » et ne deviennent des situations probabilistes qu'après construction d'un modèle d'urne [20] qui rend compte — c'est l'étape de modélisation — des résultats d'une expérience aléatoire ; le calcul permet alors d'assigner une loi de probabilité au phénomène étudié dont la pertinence sera jugée a posteriori par l'adéquation des résultats prédictifs du calcul aux résultats expérimentaux.

Dès que la notion d'expérience aléatoire — définie par un protocole expérimental — est occultée, le calcul risque de n'être qu'une syntaxe vide [21] produisant des résultats — lois de probabilités — aberrants ; les textes qui suivent illustrent d'une autre manière cette dernière assertion.

IV. — Texte III. Modélisations dans le continu : les difficultés

Bertrand donne deux exemples de modélisations dans le continu qu'il réfute. L'analyse de ces textes — solutions avancées et critiques — révèle les difficultés rencontrées

17 Se reporter, pour plus de détails, à l'article de Michel Henry paru dans Repères Irem N°51 [9]

18 Dans le cas des exemples traités par Buffon la solution avancée est plausible, l'idée de mesurer que les prévisions données par le calcul est en bonne adéquation avec l'expérience n'est pas évoquée, ce qui laisse penser que Buffon croyait à la possibilité d'un traitement a priori possible du problème.

19 Pour un développement de ce point de vue on pourra se reporter à l'article A propos de la définition de la probabilité dans Enseigner les probabilités au lycée publié par la commission inter-IREM de Statistiques et probabilités et dans Autour de la modélisation en probabilité publié par les Presses universitaires Franc Comtoise. [6]

20 Dans le cas d'un nombre fini de possibles

21 H. Poincaré, qui a longuement analysé les paradoxes de Bertrand dans Calcul des Probabilités de 1912 [5], distingue clairement entre le point de vue sémantique et le point de vue syntaxique. Un traité de calcul des probabilités envisagé comme exposé d'une théorie mathématique ne peut traiter que du point de vue syntaxique. Assigner des probabilités à des évènements , des lois de probabilité à des phénomènes n'est pas de la compétence du mathématicien mais de celle d'un expérimentateur sur la base de données statistiques. La définition complète de la probabilité est une sorte de pétition de principe : comment reconnaître que tous les cas sont également possibles ? **Une définition mathématique ici n'est pas possible.** […] Ainsi tout problème de probabilité offre deux périodes d'étude : la première métaphysique pour ainsi dire, qui légitime telle ou telle convention ; la seconde , mathématique , qui applique à ces conventions les règles du calcul.

lors des premières tentatives de modéliser des phénomènes aléatoires dans le continu. Ils permettent de mettre en évidence un deuxième type d'obstacle épistémologique rencontré lors des premières tentatives pour modéliser des phénomènes dans le continu.

1. La cible. Extrait du Calcul des Probabilité (1889) de J. Bertrand [4]

PROBLÈME XVI. — On tire à la cible. L'arme, sans être parfaite, ne présente aucun défaut systématique ; les déviations ont en tous sens même probabilité. L'hypothèse est-elle réalisable ? On le suppose.

Quelle est la probabilité pour que le point frappé soit à une distance du but comprise entre r et r+ dr ?

Les données sont insuffisantes, cela semble évident[22]. On a résolu cependant le problème par une fausse application des principes.

Rapportons le point où frappe la balle à deux axes de coordonnées ayant pour origine le centre de la cible, c'est-à-dire le point que l'on veut atteindre. Soient $Q(x)dx$ la probabilité inconnue pour que l'abscisse du point frappé tombe entre x et $x + dx$, $Q(y)dy$ la probabilité pour que l'ordonnée tombe entre y et $y + dy$. La probabilité pour que 1a balle frappe le rectangle $dxdy$, dont les coordonnées sont x et y est, d'après le principe des probabilités composées,

$$Q(x)Q(y)\,dxdy.$$

Cette probabilité, d'après notre hypothèse, ne doit dépendre que de la distance du point frappé à l'origine, et l'on doit avoir $Q(x)Q(y) = F(x^2 + y^2)$.

Cette équation suffit pour déterminer la fonction Q. On en déduit en prenant les dérivées successivement par rapport à x et à y,

$$\frac{Q'(x)}{xQ(x)} = \frac{Q'(y)}{yQ(y)}$$

La fraction $\frac{Q'(x)}{xQ(x)}$ est par conséquent constante, et l'on en conclut que la fonction Q(x) qui doit s'annuler quand x est infini, est de la forme $Q(x) = Ce^{-k^2x^2}$.

Ce résultat, fort remarquable, n'est pas, malheureusement, acceptable.

La connaissance de la valeur de x changerait, en effet, la probabilité de celle de y et le facteur par lequel il faudrait multiplier $Q(x)dx$, pour obtenir la probabilité d'un écart y dans un sens, associé à un écart x dans l'autre, serait une fonction inconnue de x et de y, très différente peut-être de $Q(y)$.

La déviation de la balle dépend, en effet, du soin plus ou moins grand et plus ou moins habile avec lequel le coup a été préparé. Si l'on a réussi sous un certain point de vue, il y a plus de chances pour que, le coup soit bon et que tous les écarts soient petits en même temps. La démonstration précédente ne tient aucun compte de cette remarque ; **les probabilités y sont traitées comme indépendantes**.

Commentaires

- Le problème n'étant pas lié à une expérience réelle est totalement indéterminé et non résoluble. La qualité de l'arme — bombardement aléatoire lorsqu'on est au-dessus de la cible, ou bombardement guidé par laser — intervient dans la valeur des paramètres de dis-

22 On a surligné en gras les phrases qui feront l'objet d'un commentaire.

persion, voire dans la forme de la loi ; on peut traiter le problème en faisant des hypothèses de modélisation [23] mais un certain nombre de paramètres resteront indéterminés [24].

• L'hypothèse de modélisation faite : (X, Y) étant le couple des coordonnées aléatoires de l'impact du tir, les événements $(x < X < x + \mathrm{d}x)$ et $(y < Y < y + \mathrm{d}y)$ sont indépendants est, comme le relève Bertrand [« *La démonstration précédente ne tient aucun compte de cette remarque; les probabilités y sont traitées comme indépendantes* »] dépourvue de toute plausibilité, comme le montre la contradiction entre le résultat produit par le calcul et celui produit par l'argument de bon sens qu'il donne.

• Il peut paraître étonnant que le problème *« Quelle est la probabilité pour que le point frappé soit à une distance du but comprise entre r et r + dr ? »* soit traité à partir des variables *(X, Y),* alors que la variable qu'il est naturel d'introduire est r distance de l'impact au centre de la cible. Le problème est alors que faute de données statistiques liées à l'expérience du tir, rien ne peut être dit sur r et donc rien ne permet de formuler des hypothèses pouvant conduire à la forme de la loi de probabilité que suit r .

2. La théorie cinétique des gaz. Extrait du Calcul des Probabilité (1889) de J. Bertrand

Dans l'article suivant Bertrand discute une loi de probabilité établie par Maxwell dans le cadre de la théorie cinétique des gaz. La démarche est la même que celle suivie dans le problème de la cible avec les mêmes hypothèses implicites a priori.

28. Un physicien justement célèbre, Maxwell, a proposé dans l'étude des gaz un raisonnement dont l'illusion est semblable.

Les molécules d'un gaz, suivant une théorie que nous n'avons pas à discuter, se meuvent en tous sens avec de grandes vitesses. Les directions sont réglées par le hasard aussi bien que les vitesses, mais toutes les vitesses ne sont pas également probables; la vitesse maxima et la vitesse moyenne varient avec la température. C'est la probabilité pour qu'une molécule ait une vitesse donnée qu'on espère découvrir, sans introduire d'autres conditions.

Soit $\mathrm{Q}(x)$ la probabilité pour que la composante parallèle à l'axe des x de la vitesse d'une molécule prise au hasard soit x. La probabilité pour que les trois composantes soient x, y, z parallèlement aux trois axes sera, d'après le théorème des probabilités composées, proportionnelle à $\mathrm{Q}(x)\mathrm{Q}(y)\mathrm{Q}(z)$.

Mais la probabilité pour qu'une molécule ait une vitesse donnée, sans que l'on indique dans quel sens, doit être une fonction de cette vitesse, puisque toutes les directions sont supposées également possibles. On doit donc avoir :

$$\mathrm{Q}(x)\mathrm{Q}(y)\mathrm{Q}(z)= F(x^2 + y^2 + z^2)$$

et cette condition suffit pour déterminer la forme de la fonction Q. On en déduit, comme en 27 :

$$Q(x) = Ge^{-k^2x^2}.$$

La démonstration n'est pas acceptable. Le principe des probabilités composées n'a pas été correctement appliqué. Si la composante de la vitesse d'une molécule parallè-

23 En explicitant les données qui y conduisent.

24 C'est ce qui est fait ici ; si on accepte l'hypothèse de modélisation — ce qui ne peut être, voir ci-dessous — le problème reste indéterminé. En effet la densité de probabilité de (X,Y) est $F(x, y) = Ae^{-k^2(x^2+y^2)}$ et :

$$\iint_P F(x, y)dxdy = \int_0^{2\pi}\int_0^{+\infty} Ae^{k^2r^2}rdrd\theta = \frac{\pi A}{k^2} = 1$$

d'où $k^2A = \pi$, la connaissance de l'un des deux paramètres est nécessaire à la détermination des solutions. L'expérience répétée est nécessaire.

lement à l'axe des x est x, la valeur de x supposée connue influe sur la probabilité pour que la vitesse composée parallèlement à l'axe des y soit y. Si, par exemple, x est égal à la vitesse maxima, le mouvement est certainement dirigé parallèlement à l'axe des x, et la probabilité de y est nulle.

La conclusion obtenue, indépendamment des objections que la théorie peut faire naître, ne mérite donc aucune confiance.

La critique de Bertrand porte, comme précédemment, sur la contradiction entre le résultat produit par le calcul et ce que permet de dire — sans calcul — la théorie physique lorsque la vitesse est maximum et dirigée selon l'un des axes du repère. Cette conclusion invalide la démarche suivie par un argument externe aux mathématiques.

Quelques conclusions

Ces fausses solutions dans le sens — *aux résultats en inadéquation avec les résultats expérimentaux* — attestent d'une difficulté à modéliser [25] dans le continu rencontrée par les pionniers. Cela révèle, pensons-nous, la manifestation de schémas de pensée profondément inscrits dans la pensée de l'époque qui se dressent en obstacle à la bonne appréhension du problème à traiter. Il s'agit, pensons-nous, d'un nouvel obstacle épistémologique que nous qualifierons de *Syndrome de Newton* [26] ; nous pouvons le caractériser de la manière suivante : la croyance que l'on peut décrire tout phénomène physique à partir de quelques postulations — qui semblent raisonnables — posées a priori. Cette croyance dérive de l'exposé de la mécanique donné par Newton dans les *Principes mathématiques de la philosophie naturelle* dans lequel tous les résultats dérivent de quelques énoncés posés sous forme d'axiomes [27].

Il faut remarquer, en dépit de la fausseté manifeste des résultats produits, en terme d'adéquation au phénomène étudié, que dans les deux problèmes évoqués par Bertrand :

1. il y a génie à penser que leur traitement peut relever du calcul des probabilités [28] et à penser qu'il faut introduire des probabilités dans le continu.
2. il y a tâche immense : comment calculer les probabilités cherchées alors qu'aucun modèle n'est disponible, aucune tradition de recherche n'est en place ?

S'il y a erreur, c'est par action inconsciente du *syndrome de Newton* qui opère par la croyance que la loi peut être obtenue à partir de postulations a priori et en conséquence par empêchement de penser la loi dans son lien avec une expérience aléatoire. La leçon de J. Bernoulli semble perdue, à savoir que l'expérience répétée est le seul moyen :

- d'avoir connaissance de la forme de la loi cherchée
- de poser les relations auxquelles elle doit satisfaire
- d'assigner des valeurs aux paramètres qui la caractérise
- de tester la bonne adéquation d'un modèle théorique aux données statistiques .

25 Ce terme anachronique sera utilisé pour éviter de longues périphrases.

26 Donner un nom évitera de nombreuses périphrases dans la suite .

27 Les axiomes pris par Newton ont une base expérimentale solide : tous les travaux faits sur la chute des corps depuis Galilée, la connaissance depuis Kepler de la trajectoire des planètes. Ses différentes postulations, hypothèses – lois fondamentales de la dynamique, loi d'inertie, loi de la gravitation – devaient ou permettre de retrouver, à l'issue des calculs conduit dans la théorie, les résultats déjà connus par observation ou expérience ou être rejetés.

28 Introduire l'aléatoire dans le monde déterministe de la physique

V. — Remarques finales

Si l'on admet les conclusions précédentes, on admettra que le premier enseignement sur les lois de probabilités continues ne peut pas être coupé de la notion d'expérience aléatoire et que le lien entre histogramme des fréquences — produit par la répétition n fois de l'expérience — et densité de probabilité est fondamental et premier [29].

L'introduction — suggérée par les commentaires du programme de TS — d'une première loi continue par l'étude de la radioactivité selon le point de vue microscopique permet de mettre ces liens en évidence. En effet, l'observation expérimentale du nombre de désintégrations d'une masse radioactive donnée dans un temps donné permet de faire observer :

- la variabilité des histogrammes obtenus pour n répétitions du comptage des désintégrations dans le temps Δt donné
- la diminution de l'ampleur de la variabilité de ces histogrammes lorsque n augmente et la quasi stabilité de ceux-ci lorsque n est grand (ordre de 10 000). Ceci suggère que le nombre $X_t(\Delta t)$ d'atomes désintégrés entre les temps t et $t + \Delta t$ est une variable aléatoire dont la loi de distribution est approchée par les histogrammes précédents.

De plus la physique fournit des hypothèses sur le comportement des atomes radioactifs qui conduisent à la loi de probabilité de la durée de vie Z d'un atome [30] qui est alors pensé comme variable aléatoire à valeur dans le *continu* $[0 ; +\infty[$. Le lien entre $X_t(\Delta t)$ et Z peut se faire à l'aide de l'étude macroscopique faite en physique [31]. Cela conduit à l'interprétation probabiliste du nombre $N(t) = N_0 e^{-\lambda t}$ et donne l'occasion de montrer comment la théorie rend compte des observations faites sur la décroissance radioactive et les fluctuations des observations à une date donnée [32].

Cette étude est l'occasion d'une véritable activité scientifique interdisciplinaire ; elle donne l'occasion de montrer les spécificités des deux disciplines physique et mathématique et la nature des liens qu'elles entretiennent. Elle donne l'occasion de montrer que les mathématiques créent des syntaxes, des cadres de pensée, des techniques de calcul pour répondre à des besoins qui leur sont parfois purement externes [33].

29 Ceci ne doit pas être perdu de vue dans l'enseignement ; l'importance du calcul des probabilités tient au fait qu'il s'applique et l'essence de cette applicabilité réside dans les monuments intellectuels que sont le théorème de J. Bernoulli et le théorème central limite. Il ne doit pas être perdu de vue que la syntaxe du calcul des probabilités s'est développée en relation avec des applications.

30 $p(Z < t) = 1 - e^{-\lambda t}$.

31 Les hypothèses formulées par les physiciens lors de l'étude microscopique du phénomène conduisent à poser que $X_t(\Delta t)$ suit une loi binomiale de paramètres N(t), p(Δt) où N(t) est le nombre d'atomes au temps t et p(Δt) est la probabilité de désintégration d'un atome dans le temps Δt . On a alors $E(X_t(\Delta t)) = N(t) - N(t + \Delta t) = N(t)(1 - e^{-\lambda t}) = N(t)p(\Delta t)$.

32 Voir article à paraître : Quelques éclairages sur la radioactivité.

33 L'exemple de la théorie cinétique des gaz de Maxwell et de la radioactivité illustrent cet aspect de l'activité mathématique.

Bibliographie

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

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

[3] LAPLACE Pierre-Simon (1749-1827) : *Théorie analytique des probabilités*, Courcier imprimeur Paris 1814.

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

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

[6] Commission inter-IREM Statistiques et probabilités, *Enseigner les probabilités au lycée*, Presses universitaires Franc Comtoise, 1997.

[7] HARTONG J., *Probabilités et statistiques. De l'intuition aux applications*, Ed. Diderot.

[8] Repères-IREM n° 32 (juillet 1998), *Spécial Probabilités*, Ed. Topiques.

[9] HENRY Michel, Des lois continues en terminale S, pourquoi et pour quoi faire ? Repères-IREM n° 51 (avril 2003), Ed. Topiques.