RAISONNEMENT EN PROBABILITÉ

*Michel **BRISSAUD***
Lycée Jean Lurçat
88600 BRUYERES)

INTRODUCTION

L'enseignement des probabilités est considéré depuis longtemps par de nombreux professeurs de lycée comme manquant de rigueur. L'apparition des arbres de probabilités dans les programmes de terminale en 1998 (utilisés comme "outils de démonstration") et celle plus récente de l'énoncé "vulgarisé" de la loi des grands nombres dans les nouveaux programmes de première ne règlent pas le malaise, au contraire.

Pourtant, il est tout à fait possible de présenter un exposé "rigoureux", mais à condition d'avoir bien conscience que dans l'étude d'un phénomène aléatoire, il y a inévitablement une phase de raisonnement qui ne ressort pas de la déduction mathématique, mais qui se place dans le cadre d'une attitude scientifique, avec sa propre rigueur.

Ceci n'est pas toujours bien compris et il en résulte des confusions qui trainent dans l'enseignement des probabilité depuis des dizaines d'années ; il est grand temps d'en sortir.

Je me place pour la suite uniquement dans le cas étudié pour l'instant en lycée, où l'univers (ensemble des résultats possibles) est fini et je présente à la fin une façon simple d'aborder certaines situations élémentaires que l'on voudrait étudier dans l'enseignement secondaire.

LA THÈSE FONDAMENTALE DES PROBABILITÉS

La première question qui se pose lors de l'étude d'un phénomène aléatoire (ou plutôt d'un phénomène que l'on veut étudier avec la théorie des probabilités) est celui du choix de la loi de probabilité sur l'ensemble des résultats possibles.

Elle peut facilement être présentée avec le célèbre problème de la punaise (tombera-t-elle ou pas sur la tête ?). Si on fait plusieurs séries d'un "grand" nombre de lancers, on constate obtenir à chaque fois des fréquences voisines, ce qui n'est pas le cas pour quelques lancers. D'où on peut raisonnablement espérer obtenir encore des fréquences voisines si on refaisait une série d'un grand nombre de lancers. Ceci n'est évidement pas démontrable. C'est un énoncé (une thèse) accepté à la lumière de résultats expérimentaux réels. Et on prend alors comme probabilité une <u>estimation prévisionnelle</u> de cette fréquence espérée.

Ce point de vue permet d'introduire dès la classe de seconde la notion théorique de probabilité, même si on ne dispose pas des règles axiomatiques de calcul. On peut ensuite présenter plus clairement la simulation de quelques expériences avec un générateur de nombres pseudo-aléatoire.

Mais comment faire cette estimation de fréquence espérée?

DES STAT AUX PROBA : UN RAISONNEMENT INDUCTIF

La démarche qui permet de passer de résultats statistiques ou de considérations physiques (pièce équilibrée...) au choix d'une loi de probabilité n'est pas dans le domaine du raisonnement déductif du mathématicien.

Il s'agit d'un raisonnement inductif : on veut "poser" des hypothèses et non en déduire des conséquences. Ce raisonnement peut être guidé par des règles acceptées par l'ensemble de la communauté scientifique (on ne fera pas n'importe quoi, on n'ira pas consulter de voyante même extra-lucide !). On se trouve là dans la position du scientifique qui devra argumenter son choix et non pas le prouver logiquement, ce qui est impossible. En retour, le modèle étant choisi, des conséquences logiques issues de la théorie devront être confrontées de nouveau (si c'est possible) à la réalité.

On trouve ici tout simplement la "démarche scientifique" à laquelle les enseignants de mathématiques ne sont pas habitués. Dans la plupart des exercices traités au lycée, le modèle mathématique est contenu dans l'énoncé (tirage au hasard, dé équilibré, donc équiprobabilité, ce "donc" ne signifiant d'ailleurs pas une déduction mathématique). Cette phase de raisonnement est alors presque systématiquement occultée.

Dans le cas où on choisit la probabilité à partir de résultats statistiques, il faut remarquer aussi qu'il n'y a pas unicité du modèle mathématique. Ainsi avec le problème de la punaise, on peut prendre comme probabilité les fréquences observées lors de 100 lancers, mais on pourrait aussi bien la lancer 1000 fois.

Ce sont d'autres constations expérimentales, plus compliquées car avec un autre niveau de paramètrage, qui montrent (mais ne démontrent pas) que le modèle sera meilleur avec 1000 plutôt qu'avec 100. Les nouveaux programmes de classe de seconde veulent aborder cet aspect à travers l'étude de la "fluctuation d'échantillonnage", sans déboucher sur la notion mathématique de probabilité qui n'est vue qu'en première.

ADÉQUATION THÉORIE - RÉALITÉ ET LOI DES GRANDS NOMBRES

Le choix des axiomes est fait évidemment pour avoir une théorie en accord avec la réalité observée, et il est facile de montrer aux élèves de lycée que la formule $P(A \cup B) = \ldots$ s'inspire d'une relation semblable entre les fréquences à travers le raisonnement du scientifique qui dégage une loi générale à partir de l'étude de la réalité.

C'est la démarche du physicien par exemple avec la loi d'Ohm : issue de constatations expérimentales répétées, elles devient ensuite "axiome" dans une théorie de l'électricité.

Cette formule P(A∪B) = ... est bien un axiome de la théorie des probabilité : on a l'occasion ici d'expliquer ces deux termes aux élèves.

Mais il est utile aussi de contrôler ensuite les développements de cette théorie, c'est-à-dire les théorèmes établis, par rapport à la réalité observée. Ainsi, le théorème dit "loi des grands nombres" est rassurant à postériori, il conforte après coup dans l'idée que les axiomes des probabilités avaient été bien choisis, car il "colle" à des observations sur les fréquences. Plus précisément, il confirme la "justesse" de la schématisation théorique de la répétition d'une expérience aléatoire.

Mais il n'apporte pas la preuve de l'adéquation théorie - réalité : ceci n'est pas démontrable. Il s'agit d'une thèse qui ne s'énonce pas dans le cadre d'une théorie déductive, on ne peut la défendre que par une argumentation.

Et ce n'est pas ce théorème des grands nombres qui nous dit qu'il faut prendre pour probabilité une fréquence observée lors de la répétition d'un grand nombre d'expériences identiques, contrairement à ce que laissent entendre les nouveaux programmes de première. Ce théorème permet de valider la théorie des probabilités, mais il ne peut en aucun cas valider des constatations expérimentales !

On pourrait donner aux élèves l'énoncé suivant : "On constate expérimentalement qu'en répétant N fois une même expérience la dispersion sur les fréquences diminue quand N augmente".

C'est cet énoncé qu'il faudrait appeler "loi des grands nombres", et on peut en commentaire dire aux élèves que les développements de la théorie permettent de démontrer des théorèmes dits "des grands nombres" dont les énoncés sont cohérents avec les constatations expérimentales et qui laissent penser que cette théorie est en adéquation avec la réalité.

De même que lorsqu'on démontre que les 3 hauteurs d'un triangle sont concourantes, on a bien un énoncé cohérent avec les dessins, avec la réalité observée, qui permet de dire que la géométrie élémentaire est bien adéquate à la réalité observée à notre échelle.

La théorie peut être vue alors comme une "explication" de la réalité, c'est une construction intellectuelle mais, pour reprendre une expression à la mode, ce n'est que du virtuel !

INDÉPENDANCE : LES CONFUSIONS CONTINUENT

Il y a depuis lontemps confusion entre la notion d'indépendance entre événements et celle d'indépendance d'expériences aléatoires.

(Suite page 8)

Prenons l'exemple classique : 2 tireurs tirent sur une cible, l'un touche avec une probabilité de 0,3, l'autre de 0,6. Quelle est la probabilité qu'ils touchent tous les deux?

La probabilité de 0,3 est sans doute obtenue après une étude statistique sur les tirs du premier, pris isolément, de même que celle de 0,6. On peut supposer de plus que le résultat d'un tireur n'influe pas sur l'autre (ce qui nécessiterait un contôle statistique ou des précautions dans le protocole expérimental) .

Pour répondre à la question, il faudrait commencer par schématiser l'expérience "couplée" ; mais la notion de probabilité produit sur un produit cartésien est hors programme au lycée. Alors on dit aux élèves : faites un arbre de probabilité et vous verrez bien qu'en faisant des produits de proba, cela marche (on trouve la réponse attendue par le prof !). Mais je n'ai encore jamais trouvé dans un manuel de lycée une justification correcte de cette règle du produit des probabilité dans cette situation.

Dire que les événements "le tireur 1 touche" et "le tireur 2 touche" sont indépendants en probabilité est une véritable escroquerie tant qu'on n'a pas modélisé l'expérience "couplée" : cet exercice est donc infaisable proprement, tout comme les exercices classiques du genre "on tire à pile ou face pour choisir l'urne dans laquelle on prend une boule" que l'on explique souvent en faisant appel aux probabilité conditionnelles avant même d'avoir une probabilité ! Un autre exemple remarquable est celui du sujet du bac 1998 (le nombre de clients prenant de l'essence ...).

Il est pourtant tout à fait possible de justifier la règle du produit des probabilité dans un arbre, à condition de comprendre qu'elle ne se démontre pas ; elle relève du choix d'un modèle, elle se justifie par une argumentation scientifique, pas par une preuve mathématique. Et c'est tout à fait faisable au niveau de la classe de première, car cela ne fait pas appel aux notions d'événements indépendants et de probabilité conditionnelle.

D'autre part, dans l'expression "répétition d'expériences identiques et indépendantes", le mot "indépendantes" est mathématiquement inutile. Si les résultats d'une expérience influaient sur l'autre, elles ne seraient pas identiques. C'est en fait ici une expression dans un discours d'argumentation scientifique (et non mathématique) relatif au protocole expérimental : il faut remettre la carte, remélanger le jeu avant de retirer... On ne peut donc pas directement la relier à la notion d'indépendance d'événements comme on le lit dans de nombreux ouvrages.

J'indique ci-dessous une façon très simple et scientifiquement rigoureuse d'enseigner cette règle du produit des probabilité dans un arbre qui permet d'étudier de nombreuses situations intéressantes dès la classe de première.

COMPOSITION D'EXPÉRIENCES ALÉATOIRES, EXPÉRIENCES INDÉPENDANTES

Il s'agit de construire un modèle mathématique en argumentant scientifiquement à partir de constatations statistiques. Je le présente d'abord à l'aide d'exemples très simples.

PRÉALABLE : RÉDUCTION D'UNE EXPÉRIENCE À UN SYSTÈME COMPLET D'ÉVÉNEMENTS, REPRÉSENTATION EN ARBRE

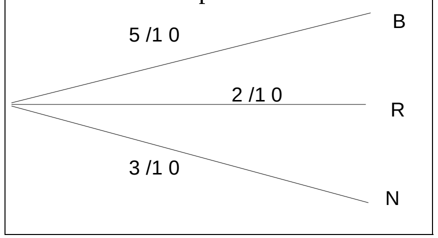
Dans une urne on a 5 boules blanches, 2 rouges et 3 noires. On tire au hasard une boule.

Un résultat possible est l'une des 10 boules. On les suppose équiprobables ("tirage au hasard"), les probabilité seront égales aux proportions de cas favorables par rapport aux cas possibles.

On note B : "elle est blanche", de même R et N. Les 3 événements B, R, N forment un système complet d'événements.

On peut alors réduire cette expérience à l'univers {B, R, N} avec les probabilité p(B) = 5/10, ...

Arbre de probabilités :



COMPOSITION CONDITIONNELLE DE 2 EXPÉRIENCES

Étude d'un exemple d'école :

Expérience 1 : on lance un dé à 4 faces numérotées (dé tétraèdrique régulier, non truqué).

Expérience 2 : on tire une boule dans l'urne décrite ci-dessus.

Règle de composition : on lance le dé ; si on obtient 3, on tire dans l'urne, sinon on ne fait rien.

Quelle est la probabilité d'obtenir une boule blanche ?

Il faut d'abord schématiser l'expérience composée. Dessiner un arbre permet d'illustrer les différentes issues, mais comment calculer les probabilité ?

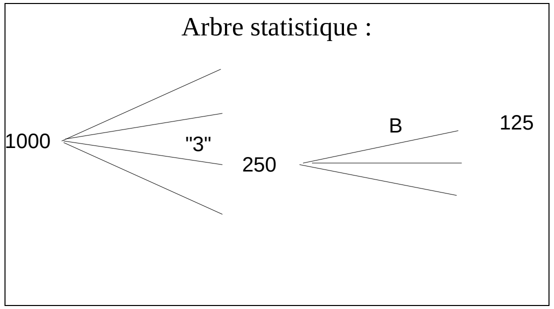
On peut prendre comme univers {1, 2, (3,B), (3,R), (3,N), 4} (c'est un système

(Suite page 10)

complet d'événements).

Approche statistique : imaginons qu'on réalise un grand nombre de fois cette expérience composée, par exemple 1000 fois.

Dans environ 25% des cas, on obtiendra un "1".



De même, dans environ 25% des cas, soit environ 250 fois, on aura un "3".

Dans ces 250 cas, on tire une boule, et on aura une blanche environ une fois sur deux (la probabilité de tirer une blanche dans cette urne est de 0,5).

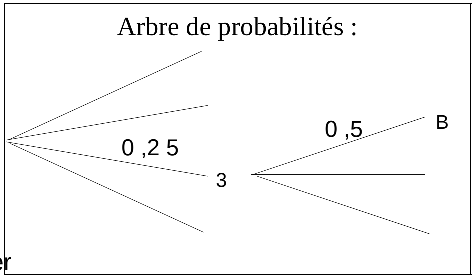
On aura alors une blanche environ 125 fois, ce qui fait une fréquence dans l'expérience composée de 125/1000.

Cette fréquence peut s'obtenir aussi par le produit 0,25 x 0,5 (fréquence du 3 pour le dé x fréquence de B pour l'urne).

CHOIX D'UNE PROBABILITÉ : il est scientifiquement raisonnable de prendre une probabilité calculée avec une règle semblable à celle observée sur les fréquences.

On prendra alors pour l'expérience composée une probabilité p telle que :

$p(3, B) = p_1(3) \times p_2(B)$ en notant p_1 la probabilité relative au dé et p_2 celle relative à l'urne, et $p(1) = p_1(1)$...



Remarques : On peut facilement étendre cette démarche et considérer la composition de plusieurs expériences. Cela permet d'aborder les tirages sans remise.

On justifie ainsi la représentation en arbre avec la règle du "produit sur un chemin", sans faire appel à la notion de probabilité conditionnelle qui ne pourrait être introduite qu'après avoir défini la probabilité sur l'univers.

COMPOSITION DE 2 EXPÉRIENCES INDÉPENDANTES

Exemple : 2 tireurs tirent sur une cible, l'un touche avec une probabilité de 0,3, l'autre de 0,6. Quelle est la probabilité qu'ils touchent tous les deux?

Vu l'énoncé, on suposera que la probabilité de toucher pour un tireur ne dépend pas du résultat de l'autre (sinon, il y aurait des indications dans l'énoncé !). Ceci sera traduit par une règle de composition qu'on peut qualifier ici d'inconditionnelle : le premier tireur tire et qu'il touche ou pas, le 2° tire aussi avec la même probabilité de 0,6 pour toucher. C'est dans cette situation qu'on peut parler d'expériences indépendantes (en un sens "mathématique" et plus seulement au niveau d'une description expérimentale).

Pour le premier tireur : T1 pour "il touche" et NT1 pour l'événement contraire. p_1(T1) = 0,3.

Pour le deuxième : T2 pour "il touche" et NT2 pour l'événement contraire. p_2(T2) = 0,6.

La démarche énoncée ci-dessus sur le choix d'une probabilité reste évidemment valable, car on est dans une situation de composition d'expériences.

L'expérience composée se schématise alors par :
{(T1,T2), (T1,NT2), (NT1,T2), (NT1,NT2)} avec p(T1,T2) = p_1(T1)· p_2(T2)...

Remarques : dans une représentation en arbre, l'indépendance des expériences se traduit par l'identité des 2 sous-arbres de 2° n iveau. On retrouve la notion de probabilité produit sur un produit cartésien, qui permetrait une autre définition mathématique de la notion d'indépendance d'expériences, mais il n'est pas indispensable de l'évoquer au lycée.

RÉPÉTITION D'EXPÉRIENCES IDENTIQUES

Ce n'est qu'un cas particulier du précédent où on compose des expériences identiques (donc "indépendantes").

Ceci est donc abordable en classe de première sans se limiter à des cas d'équirépartition.

En classe de terminale, des considérations de combinatoire amènent facilement à la distribution binomiale.

ENONCÉ PLUS GÉNÉRAL :

Soit E1 une expérience réduite à {A1, ..., Ai, ...An} avec une probabilité p_1, et E2 une expérience réduite à {B1, ..., Bj, ...Bm} avec une probabilité p_2.

On considère la règle de composition suivante : On fait E1 ; pour un i fixé, si Ai est réalisé, on fait E2, sinon, on ne fait rien.

L'expérience composée pourra alors être représentée par {A1, ..., Ai-1, (Ai, B1), ... (Ai, Bm), Ai+1, ..., An} avec la probabilité p définie par :

pour tout k différent de i, p(Ak) = p_1(Ak)

pour tout j, p(Ai, Bj) = p_1(Ai)· p_2(Bj)

(Suite page 12)

REMARQUES :

1) La démarche adoptée pour choisir la probabilité est la même que celle qui permet de poser en axiome la formule P(A u B) = ... à partir d'une relation observée sur les fréquences.

Une étude statistique laisse penser que c'est alors une "bonne" schématisation de l'expérience composée. Ceci n'est évidemment pas démontrable.

2) On pourrait être choqué par l'univers choisi : ses éléments ne semblent pas, en première lecture, de la même nature. Mais on peut les voir comme des listes de longueur 1 ou 2, et dans un tel ensemble, il y a évidemment possibilité de faire des réunions ou des intersections et de revenir au langage ensembliste ; mais le vocabulaire des événements est beaucoup plus commode. Les profs de math n'ont malheureusement pas l'habitude de manipuler de tels ensembles contrairement aux informaticiens qui ont souvent besoin d'utiliser des "structures de données" bien plus sophistiquées.

3) (Ai, Bj) pourrait être écrit "Ai suivi de Bj", à la rigueur "Ai et Bj", mais cela n'aurait aucun sens d'écrire une intersection entre Ai et Bj. Il ne faut pas voir la formule $p(Ai,Bj) = p_1(Ai)\cdot p_2(Bj)$ comme celle qui exprime l'indépendance probabiliste de deux événements $p(A\cup B) = p(A)\,p(B)$: ça n'a (presque) rien à voir.

CONCLUSION

Il serait donc tout à fait possible d'ajuster l'enseignement des probabilités ainsi :

1) en classe de seconde, introduire la notion de probabilité à partir d'études statistiques, pour ensuite aborder la simulation à l'aide de générateur de nombres pseudo-aléatoires.

2) en classe de première :

- compléter les études statistiques de seconde pour amener un énoncé d'une "loi des grands nombres" résultant de considérations expérimentales (laisser tomber l'énoncé vulgarisé du théorème des grands nombres qui est dans les programmes actuels).
- donner (en justifiant statistiquement) les règles de calcul des proba.
- présenter la composition d'expériences aléatoires, introduire l'indépendance d'expériences, la schématisation de la répétition d'une expérience.

3) en terminale, on peut introduire (à partir de résultats statistiques) la notion de probabilité conditionnelle, l'indépendance en probabilité d'événements et étudier la loi binomiale.

(Suite page 13)

(Suite de la page 12)

N.D.L.R. On peut contacter M. BRISSAUD : Michel.Brissaud@ac-nancy-metz.fr, et consulter son site personnel :
http://perso.wanadoo.fr/michel.brissaud/math/sommaire-math.html