PUBLICATIONS MATHÉMATIQUES ET INFORMATIQUES DE RENNES

ANDRÉ TOTOHASINA

Quelques misconceptions qui risquent de devenir obstacles épistémologiques sur la notion de probabilité conditionnelle

Publications de l'Institut de recherche mathématiques de Rennes, 1993, fascicule 3
« Fascicule de didactique des mathématiques », , p. 1-34
<http://www.numdam.org/item?id=PSMIR_1993___3_1_0>

QUELQUES MISCONCEPTIONS QUI RISQUENT DE DEVENIR OBSTACLES EPISTEMOLOGIQUES SUR LA NOTION DE PROBABILITE CONDITIONNELLE

André TOTOHASINA
IRMAR — Université de Rennes 1

I Introduction

Si le concept des probabilités a vu le jour vers le milieu du 17ième siècle, l'histoire atteste que celui de probabilité conditionnelle a une apparition relativement récente. Il a fallu attendre 33 années après le début du 20ième siècle, grâce aux travaux de Kolmogorov qui a institutionnalisé les probabilités comme objets mathématiques formels, pour que la notion de probabilité conditionnelle soit doté d'objet mathématique détaché de tout contexte concret. Elle apparaît d'ailleurs comme un des moteurs du développement de la mathématique des probabilités : toute probabilité est en fait conditionnelle. Rappelons que d'un point de vue philosophique, trois grands courants d'idées sont reconnus comme critères de classement des divers spécialistes des probabilités (cf.[34], pp.1-14 chap.III -IX) :

- l'approche l*ogiciste* (ou *nécessariste,* terminologie de L.J.Savage) selon laquelle les probabilités sont une extension de la logique, en représentant le degré de confirmation d'une proposition étant donné toutes les propositions auxquelles elle est liée ;
- l'approche *subjectiviste (*ou *personnaliste*, terminologie de L.J.Savage) qui conçoit le calcul des probabilités comme une mesure de la confiance qu'on a dans la vérité d'une proposition ; on checrhe à utiliser le calcul des probabilités dans la plupart des situations d'incertitude, y compris des situations subjectives ;
- l'approche *objectiviste* ou *fréquentiste* selon laquelle le concept de probabilités s'applique uniquement aux événements aléatoires susceptibles de se produire lors d'épreuves répétées et indépendantes ; on ne peut probabiliser que les événements dont l'observation des fréquences d'apparition, au cours d'épreuves répétées, constitue une approche empirique. Elle est certainement d'un grand intérêt didactique, car elle autorise de véritables expérimentations, par la voie d'étude statistique, si l'on vise l'éveil du jugement probabiliste chez les enfants. Beaucoup de travaux de recherche sont déjà engagés sur cette voie [19], [6], [5], [13], [14], [16] etc. Ils concernent pour la plupart le concept de probabilités en général.

Le présent article relate essentiellement la notion de probabilité conditionnelle, discipline d'enseignement. Une enquête que j'ai menée auprès d'étudiants de Deug en 1991 a montré l'existence de deux conceptions réductrices relatives au concept de probabilité conditionnelle [34], les conceptions causaliste et chronologiste, sur lesquelles peuvent se fonder de véritables obstacles épistémologiques [1], [8] et [9]:

- <u>conception chronologiste</u> d'une probabilité conditionnelle : c'est le fait de sous-tendre systématiquement de "chronologie " ou de "temps" la façon d'appréhender ce concept ;
- <u>conception causaliste</u> d'une probabilité conditionnelle : c'est le fait d'introduire (implicitement) ,une relation de "cause à effets" ou de "cause à conséquences" dans la façon

d'appréhender ce concept.

Se pose alors la question de savoir en quoi justement ces conceptions risquent de présenter des obstacles dans la façon de résoudre des problèmes qui relèvent de la notion de probabilité conditionnelle. Il semble que la *réversibilité* de la probabilité conditionnelle ne soit pas possible si les conceptions chronologiste et causaliste demeurent prégnantes chez les élèves. En effet, la conception chronologiste suppose que dans l'expression "la probabilité de B, sachant A", A précède B, ou A préexiste avant B; par suite il n'est pas possible d'exprimer la "probabilité de A, sachant B".

La conception causaliste suppose que dans l'expression "la probabilité de B, sachant A" A est la cause de B, donc c'est un non sens d'exprimer la "probabilité de A, sachant B".

De plus, il apparaît également que le mode traditionnel qui consiste à introduire cette notion par le rapport de cardinaux engendre une autre conception fausse, la conception cardinaliste[35] : c'est le fait de transporter systématiquement la quantification proportionnelle d'une probabilité correspondant à une hypothèse d'équiprobabilité d'événements élémentaires "cas favorables sur cas possibles" à la probabilité conditionnelle ; elle se concrétise par la tendance systématique à se représenter la probabilité conditionnelle de A sachant B, $p_A(B)$, par le rapport $p_A(B) = \frac{\text{Card}(B \cap A)}{\text{CardA}}$ ou , à tort, par $p_A(B) = \frac{\text{CardB}}{\text{CardA}}$; elle s'<u>oppose à la conception fondamentaliste</u> qui consiste à utiliser la définition formelle de ce concept.

A mon sens, les deux premières conceptions relèvent du cognitif, et la troisième de l'opératoire. Ce dernier peut d'ailleurs s'expliquer par le phénomène de plongement, pour reprendre la terminologie de F.PLUVINAGE[30'], à cause de la situation d'introduction, donc plongement d'origine didactique.

Dans le présent article, je propose une ingéniérie qui tente justement de contrôler ces trois représentations de cette notion.

II Objectifs

Mon objectif est double :

Primo : essayer de donner à la notion de conditionnement d'un événement et à celle de probabilité conditionnelle un sens intuitif. En effet, l'incompréhension de cette notion de conditionnement d'événement me semble une des origines des obstacles qui seraient en amont de certaines difficultés récurrentes, telles la *confusion entre* $p(A \cap B)$ *et* $p_B(A)$, *la* confusion entre la notion d'*indépendance*[2] *stochastique* et celle d'*incompatibilité* d'événements, le comportement *cardinaliste,* etc.

Par contre, la formule de probabilités composées peut apparaître d'une manière presque spontanée, même chez des élèves non initiés à cette notion de probabilité conditionnelle.

Secondo : introduire la notion en question à l'aide d'une résolution de problème , un

[2] Il conviendrait de ne pas utiliser ce terme isolément, pour faire déjà une distinction avec d'autres types d'indépendance rencontrés également en mathématiques, en l'occurrence *l'indépendance linéaire* etc, ou ailleurs.

problème qui mettrait en œuvre le théorème de Bayes (d'une manière implicite) sans que l'élève ait besoin de mobiliser les prérequis supposés nécessaires (entre autres, nous pensons à l'acquisition de la théorie élémentaire des ensembles : intersection, réunion etc). Inutile de rappeler que ces derniers sont absents du programme scolaire des mathématiques en vigueur dans le secondaire. Alors comment faire? Ma première réponse s'oriente vers l'utilisation d'un autre registre, le registre graphique (en l'occurrence un arbre des probabilités), en plus du registre numérique. On tentera donc de partir d'une rationalité vécue quotidiennement, qui a en fait la signification d'une probabilité, en l'occurrence la notion de pourcentage (de chance). Remarquons que cette dernière met souvent *en acte* la notion de probabilité.

III Hypothèse : attente sur les réactions d'élèves

Remarque historique préliminaire

Dans l'Encyclopédie Méthodique des Sciences (article PROBABILITE, p.644 : un texte qui date du XIXème siècle), nous apprenons qu'à l'époque, du moins en France, la notion de probabilité conditionnelle n'avait pas encore le statut de concept formalisé. Mais, comme l'extrait suivant va nous le montrer, elle était déjà bien présente implicitement : elle était pratiquée à l'intérieur de la notion de *probabilité composée*. Cette dernière était définie comme suit (Sic):

> *"Une probabilité composée est la probabilité d'un événement qui ne peut arriver qu'au cas qu'un autre événement , lui-même finement probable, arrive".*
> *Un exemple va l'expliquer.*
> *Je suppose que dans un jeu de quadrille de 40 cartes, l'on me demande de tirer un cœur, la probabilité de réussir est 1/4 de la certitude, puisqu'il y a 4 couleurs et 10 cartes de chaque couleur également possibles. Mais, si l'on me dit ensuite que je gagnerai si j'amène le roi de cœur, alors la probabilité devient composée ; car,*
> *1° :il faut tirer un cœur, et la probabilité est 1/8 ; 2° : supposé que j'ai tiré un cœur, la probabilité sera 1/10, puisqu'il y a 9 autres cœurs que je peux aussi bien tirer que le roi.*
> *Cette probabilité entre sur la première, n'est que la dixième d'un quart, ou le 1/4 de 1/10, c'est-à-dire 1/40 de la certitude. Cette probabilité composée s'estime donc en prenant de la première une partie, telle qu'on la prendroit de la certitude entière, si cette probabilité étoit une certitude.... On peut appliquer ce calcul à toute sorte de preuves ou de raisonnements, réduits pour plus de clarté à la forme prescrite par l'art de raisonner : si l'une des prémices est certaine & l'autre probable, la conclusion aura le même degré de probabilité que cette prémice; mais si l'une et l'autre sont simplement probables, la conclusion n'aura qu'une probabilité de probabilité, qui se mesure en prenant de la probabilité de la majeure, une partie telle que l'exprime la fraction, qui mesure la probabilité de la mineure."*

Il est clair que l'on parle bien (intuitivement certes) d'une probabilité conditionnelle dans ce 2° malgré son absence conceptuelle, contrairement à la *probabilité composée*, où il y a composition d'événements. Soit : en utilisant le symbolisme moderne, en désignant par C l'événement "obtenir un cœur", par R l'événement "obtenir un roi", la locution employée au 2°) "*supposé que j'ai tiré un cœur*" exprime un conditionnement qui est aléatoire ; la formule instanciée est $p(C \cap R) = p_C(R) \times p(C)$: *la quantité* $p_C(R)$ est acceptée intuitivement comme une probabilité, p_C étant une probabilité effective, ayant les mêmes propriétés formelles que p. Cela correspond à ce que G.VERGNAUD[37] appelle un *théorème-en-acte* [1]: p_C était à l'époque admise comme probabilité, mais essentiellement par l'acte. Par ailleurs, la locution comme " *1/4 de 1/10*", qui rappelle la notion de partage, pour exprimer $1/4 \times 1/10$ paraît très intéressante également ; d'autant plus qu'une expression de ce type figurait déjà dans quelques copies d'élèves de l'année 1989/90. Elle ensemence, me semble-t-il, un champ intuitif important, comme le dit R.GRAS dans sa thèse ([19], p.21). D'ailleurs certains auteurs, comme B.V. Gnedenko & A.L.Khintchine[18], T.H.WONNACOTT et R.J.WONNACOTT[40], n'hésitent pas à affirmer que la notion de partage serait justement l'arrière-fond de la formule des probabilités composées $P(A \text{ et } B) = P(B|A).P(A)$ (lisible sous la forme : *P(B/A) de P(A)*) (les nombres étant exprimés sous forme fractionnaire). C'est donc une possibilité de *mettre* une probabilité conditionnelle *en acte*. Bien sûr, cette opération de proportion se connote également de cardinalité, mais c'est l'esprit probabiliste qui prédomine. Ce texte explique en partie le choix du mot "composée" : on parle de *la probabilité d'une probabilité.*

Choix didactiques et les hypothèses sous-jacentes

Eu égard à cette remarque et aux objectifs ci-dessus, je suis amené à conjecturer que :

- **a)** la notion de pourcentage et celle de partage faciliteraient l'appréhension de la formule des probabilités composées. Cette dernière servirait ensuite de tremplin pour faire émerger le concept de probabilité conditionnelle. On sait en effet que la notion de pourcentage possède au moins les caractères épistémologiques suivants :
  - être indissociable du *concept de proportion* ;
  - en plus, être attaché à des contextes qui évoquent le *modèle exclusivement multiplicatif*, ou tout simplement un *opérateur de multiplication.*

Mais, ceci n'est pas sans danger. Car il n'est pas rare d'entendre des propos comme "... Dans un pays X , il y a 150% d'inflation !..." . Or 150% est lu aussi en terme de fraction 150/100, soit un nombre supérieur à 1. Heureusement, pour des cas pareils, le mot *chance* est hors de question, donc cela ne doit pas nuire à l'intuition probabiliste.

On lira provisoirement, P(A) = pourcentage de chance de réalisation de A. On terminera la

---

[1] Il s'agit d'un théorème qui n'est pas enseigné aux enfants. Pourtant on le voit fonctionner spontanément dans leurs raisonnements, pour certaines valeurs numériques simples, et pour des domaines familiers de l'expérience. Il y a là une prémisse de la notion en question, en l'occurrence ici celle de probabilité conditionnelle.

séance d'introduction en faisant une courte synthèse résumant et "généralisant" les résultats obtenus.

De la relation P(A et B) = P(B/A).P(A) on tire $P(B/A) = \frac{P(A\ et\ B)}{P(A)}$. Cette étape paraît nécessaire, à mon avis, car nous savons bien qu'un concept ne se construit pas en une seule fois, il s'affine progressivement. De plus, quelquefois, certaines conceptions fausses ou approximatives peuvent également être didactiquement utiles, surtout si elles apparaissent ou sont débusquées pendant la séance du cours. Je pense qu'il est donc préférable de procéder avec des représentations qui évoluent ou qui se modifient , plutôt que de penser radicalement éliminer les conceptions fausses en leur substituant les conceptions visées. En outre, pourrait-on ainsi diminuer l'éventuelle robustesse du *comportement cardinaliste sur la probabilité conditionnelle?*

- **b)** l'analyse arborescente (c'est-à-dire l'utilisation d'un arbre des probabilités) permettrait de sauter sans trop de risque l'étape des opérations ensemblistes, comme déjà signalé plus haut, en l'occurrence la distributivité de l'*intersection* par rapport à l'*union*. De telles opérations, comme nous le savons tous, ne figurent plus au programme dans le secondaire actuel. En outre, la pratique de cette analyse arborescente pourrait diminuer également l'obstacle éventuel de *renversement de situations* du cas bayesien, grâce à ce travail dans un cadre graphique. Et la séquentialité d'une telle analyse ne ferait qu'accroître son effet utile pour résoudre un problème de probabilité conditionnelle, même celui du type bayesien qui est déjà relativement complexe.

A titre d'exemple anecdotique, mais didactiquement instructif à mon avis, je me permets de présenter la situation-problème suivante au sujet duquel un stagiaire est venu discuter :

Deux joueurs A et B jouent chacun à leur tour jusqu'à ce que le premier d'entre eux gagne.
A joue le premier et peut gagner avec la probabilité p_1 et perdre avec la probabilité $1-p_1$.
B joue ensuite et peut gagner avec la probabilité p_2...
Calculer la probabilité des événements {A gagne} et {B gagne}, en fonction de p_1 , p_2.

Ma première réaction était de lui présenter ainsi une solution :

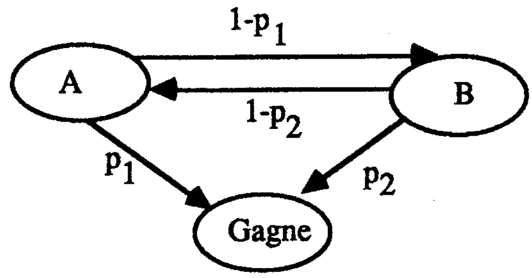
"Il s'agit d'un problème où l'espace probabilisé est infini dénombrable...A joue au 2k+1ème coup, et B au 2k ème coup... Donc les probabilités demandées sont :

$$p(\{A \text{ gagne}\}) = \sum_{k=0}^{+\infty} p(\{A \text{ gagne au 2k+1ème coup}\}) = p_1.\sum_{k=0}^{+\infty}(1-p_1)^k(1-p_2)^k = \frac{p_1}{1-(1-p_1)(1-p_2)};$$

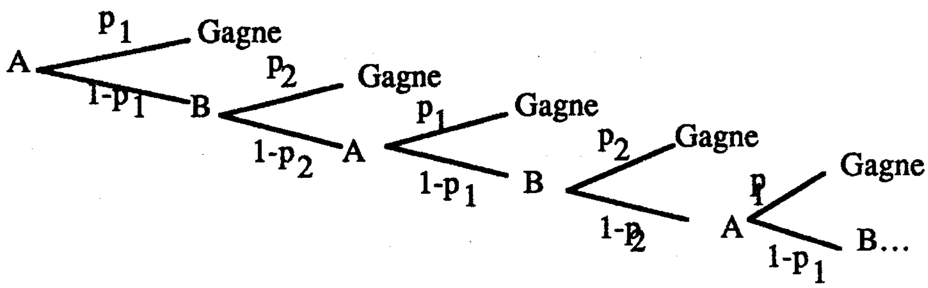
$$p(\{B \text{ gagne}\}) = \sum_{k=0}^{+\infty} p(\{B \text{ gagne au 2k ème coup}\}) = p_2.(1-p_1).\sum_{k=0}^{+\infty}(1-p_1)^k(1-p_2)^k = \frac{(1-p_1)p_2}{1-(1-p_1)(1-p_2)}".$$

Le stagiaire réagit :"Non, je ne comprends pas là..."

Alors j'ai produit le graphe ci-dessous :



Ensuite, j'ai présenté l'arbre des probabilités pour une partie à 2 reprises :



"... Durant les 9 premiers coups, on a :
{A gagne au 1er coup}, ou {A perd au 1er coup} et {B perd au 2ème coup} et {A gagne au 3ème coup}, ou ..., ou {A gagne aux 1er, 3è, 5è et 7è coups} et {B perd aux 2è, 4è, 6è et 8è coups} ..."
Enfin, le stagiaire déclara satisfait : "...Ah! oui...là, c'est plus clair...et je pense que c'est avec cette dernière figure que j'arriverais à mieux convaincre les étudiants..."

Toutefois, nous verrons plus loin une éventuelle frontière de validité du procédé par arborescence [29].

A propos de cette procédure par dichotomie arborescente, on sait que, même en s'adressant à des étudiants déjà initiés à la notion de probabilité conditionnelle dans l'enseignement supérieur, il n'est pas rare de fonder son raisonnement sur un graphe. Par exemple, dans les cas d'une chaîne de Markov finie, la classification des états peut être expliquée concrètement sur le graphe topologique correspondant (connexité d'un graphe, ou des sous-graphes,...etc). En plus de sa *richesse* en information, cela permetrait aussi de gagner énormément de temps par rapport à d'autres méthodes algébriques. Enfin, une fois cette méthode, si l'on peut dire, "maîtrisée" dans le cas de deux "conditions" et deux "produits", il serait envisageable de généraliser à 3, 4, etc.

IV Méthodologie sur l'action didactique effective

Je propose donc de :
* présenter une situation-problème qui ne contient que des nombres en pourcentage et qui ne présuppose pas l'utilisation d'une probabilité uniforme, afin d'éviter surtout une éventuelle complication d'ordre combinatoire, qui prête peu à une causalité et à la chronologie, et d'une manière évidente, à un comportement cardinaliste,

* faire émerger, faire définir l'algorithme de calcul à partir d'un exemple[17],
* donner enfin une définition formelle d'une probabilité conditionnelle,
* ne pas insister sur les opérations ensemblistes.

Notons que le but d'une telle entreprise n'est pas exclusivement de faire apprendre à résoudre un problème particulier, analogue aux techniques générales de type "problem solving" (méthodes générales de résolution de problèmes explicitées, entre autres, par des organigrammes) ou à des apprentissages algorithmisés vers lesquels se précipitent volontiers les élèves en difficulté, mais de faire comprendre à travers la méthode engagée une méthode pour traiter une classe de situations-problèmes plus ou moins isomorphes à celle présentée. On vise également à transmettre aux élèves "la technique d'abstraction" sous-jacente, et ce, en sollicitant leur participation active. Cette dernière aurait une probabilité assez forte pour faciliter l'intériorisation, c'est-à-dire l'appropriation par une majorité d'élèves, de la notion en question.

V Un exemple de situation-problème

V- A : Enoncé

Un magasin stocke un certain produit dans des boîtes. Ces boîtes sont de deux couleurs : rouge dans la proportion 25%, bleue dans la proportion 75%. Elles sont protégées par des cartons identiques entre eux. Chaque carton ne contient qu'une seule boîte. Certains cartons portent, en-dessous et à l'extérieur, la marque M , les autres ne portent aucune marque. On précise d'autre part que :

parmi les cartons contenant une boîte rouge, 45% portent la marque M ; parmi ceux qui contiennent une boîte bleue, 60% portent la marque M.

On prend au hasard un carton du magasin.

Questions :

Q1- Définir un univers des possibles.

Q2- On ouvre le carton tiré. On remarque qu'il contient une boîte rouge.
Quelle est la probabilité que le carton porte la marque M?

(Q2bis : si la boîte contenue dans le carton était bleue, quelle serait la probabilité que le carton porte la marque M?) (question à mettre en réserve).

Q3-Quel est le pourcentage des cartons qui portent la marque M?
En déduire la probabilité qu'un carton tiré porte la marque M.

Q4 - On n'ouvre pas le carton tiré. On remarque toutefois qu'il porte la marque M.
Quelle est la probabilité que ce carton marqué M contienne une boîte rouge?

On adoptera les notations suivantes pour les événements (par souci d'uniformité) :

M : "obtenir un carton marqué M". R : "obtenir un carton contenant une boître rouge".

B : "obtenir un carton contenant une boîte bleue".

V- B Prévision du déroulement de la séquence didactique

V-B.1 Plan

* Présenter le problème, sans avoir mis le titre du chapitre, objet de l'étude, "Probabilité conditionnelle". Cela éviterait une didactique trop directive. On sollicite l'élève (ou la classe) à conclure par lui-même (elle-même), après avoir mis à l'épreuve ses éventuelles conceptions. Ces dernières auraient ainsi une grande chance d'être corrigées dans le cas où elles étaient erronées.

* Les questions seront posées au fur et à mesure, et non d'un seul coup.

* Parler de l'objet du chapitre, seulement après la résolution collective de la question Q2.

* La définition axiomatique d'une probabilité conditionnelle sera donnée seulement après une petite synthèse du problème traité.

* Prévoir un test après le cours.

V-B.2 Résolution collective de la situation-problème

- **Pour la question Q1**

La classe sera invitée à proposer une réponse.

La réponse attendue est que l'univers associé se définit par : $\Omega = \{$cartons dans le magasin$\}$.
On peut s'attendre également à d'autres propositions, par exemple $\Omega = \{$bleue, rouge, marque M$\}$ qui n'est pas pertinente évidemment, car les événements "contenir une boîte rouge" et "porter une marque M" ne sont pas incompatibles.
On rappellera ensuite la définition d'une probabilité p en tant qu'application de $\mathcal{P}(\Omega)$ vers [0, 1] avec les deux axiomes classiques.

Dans l'espoir de rendre les élèves conscients de l'acception probabiliste de la notion de pourcentage ici, on demandera également les probabilités respectives de A, B et M. On pourrait ainsi s'attendre à une réaction directe des élèves à la simple lecture de l'énoncé pour les deux premiers événements. Par contre, la valeur de p(M) n'est pas immédiate.
Remarque : D'une manière plus formelle, remarquons que cette situation peut très bien être transposée sous forme d'un schéma d'urnes. Il suffirait de considérer deux urnes non équiprobables d'être tirées, soit une *urne rouge* U_r et une *urne bleue* U_b, par exemple, telles que l'urne rouge ait la probabilité égale à 0.25 d'être tirée et la bleue 0.75. Chaque urne contiendrait des boules identiques indiscernables au toucher dans les proportions suivantes : 45% des boules contenues dans l'urne rouge porteraient cette marque, et 60% des boules dans l'urne bleue portent également la marque M. L'expérience consisterait à tirer une urne, puis de cette urne extraire une boule.
A un niveau scolaire un peu plus élevé, il serait envisageable également de prendre l'espace probabilisé (Ω, β, P) avec : - comme univers $\Omega = \Omega_1 \times \Omega_2$, où $\Omega_1 = \{R, B\}$ et $\Omega_2 = \{M, \bar{M}\}$:
- comme tribu $\beta = \mathcal{P}(\Omega_1) \times \mathcal{P}(\Omega_2)$,
- comme probabilité $P = p1 \otimes p2$, où pi est la probabilité sur $(\Omega_i, \mathcal{P}(\Omega_i))$, i=1, 2.

Le problème est que dans ce cas, les événements élémentaires cessant d'être équiprobables., l'intuition probabiliste en serait affectée.

De plus, avec un tel modèle d'urnes, la chronologie appuyée des événements pourrait facilement créer une résistance chez des élèves débutants relativement à la *notion nouvelle* de probabilité conditionnelle (cf.l'introduction).

Pour la question Q2

Il s'agit de reconfirmer la position du problème, notamment l'existence et le rôle d'une *information supplémentaire*. Par exemple, on pourra poser la question comparative suivante : cette question Q2 est-elle "identique à " :

On prend au hasard un carton du magasin. Quelle est la probabilité que le carton porte une marque M?

La réponse attendue à cette question est *Non.*

En fait, il s'agit donc de calculer la probabilité de l'événement M, sous une condition bien précise, à savoir *le carton contient une boîte rouge.*

On peut encore poser une autre question pour amener l'élève à prendre conscience de l'importance de l'*information supplémentaire* (c'est-à-dire la réalisation de R ici) qui conditionne l'événement M . Elle est de type métamathématique traduisant en quelque sorte les notions de *restriction* et de *trace* mathématiques. Par exemple :

Intuitivement, cette probabilité de M qu'on se propose de calculer est-elle la même que p(M) (c'est-à-dire la probabilité de M, sans tenir compte d'aucune information supplémentaire)?

Réponse attendue : *Non.*

A propos de la notation

Cette probabilité (qui est conditionnelle) ne pourra pas être notée p(M). Comment faire pour la distinguer de p(M)?

Réponse attendue :

On recueillera les éventuelles propositions des élèves.

Il est probable que la notation indicielle $P_R(M)$ figure parmi la liste proposée. En tout cas, le problème est ainsi *dévolu* à la classe qui doit vivre une sorte d'amorce d'un *débat scientifique* au sens de Marc LEGRAND (le mot est certainement trop fort ici pour la seule activité de désignation, mais la situation est comparable à un débat car il y aura partage des points de vue, contestation, rejet, consensus,...).

On dira à la classe qu'on va noter cette probabilité (conditionnelle) d'une manière qui contienne les deux informations R et M, et qui laisse des places non symétriques aux variables considérées R et M . Une notation comme P'(M), par exemple, est à rejeter car elle ne comporte rien qui puisse rappeler R, événement conditionnant.

On demandera ensuite d'évaluer la probabilité conditionnelle $p_R(M)$.

Erreur attendue : calcul de $p(M \cap R)$ au lieu de $p_R(M)$.

Il serait profitable de signaler une telle confusion pour le reste de la classe, afin qu'elle ait quelque chance de ne plus être répétée. Mais les raisons de l'erreur peuvent être diverses. En particulier, les élèves peuvent être troublés par le fait qu' il n'y a pas de calcul à faire ici. Ce qui est peu habituel dans le contrat didactique. Il est surtout question de percevoir le rôle du conditionnement de l'événement M.

Enfin, remarquons également que la notion de temporalité (ou de chronologie) reste présente dans cette activité : on peut voir d'abord la marque, puis après la couleur de la boîte. Par contre la notion de causalité est pratiquement effacée, sauf celle inférée par l'ordre dans lequel le texte est donné.

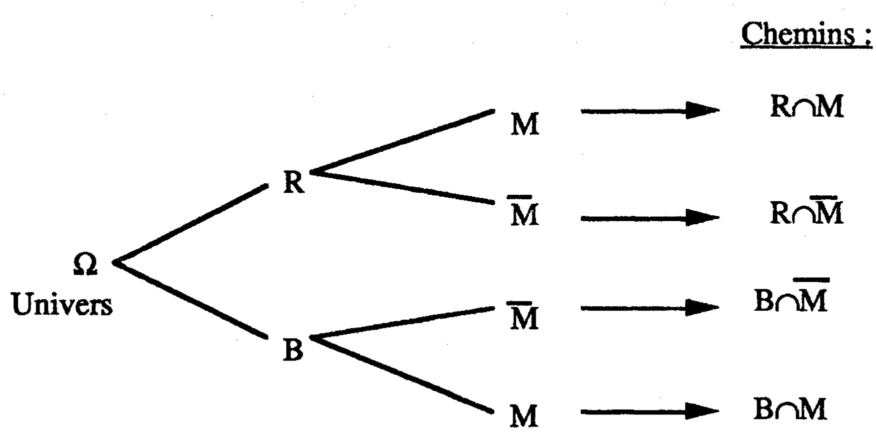
A propos du travail sur le cadre graphique [10] : arbre des probabilités

On peut faire allusion à un "arbre de dénombrement ", technique traditionnellement déjà pratiquée dans le chapitre sur le dénombrement, antérieur au présent chapitre des probabilités. On pourra demander aux élèves de s'en inspirer en cas d'éventuelle hésitation.

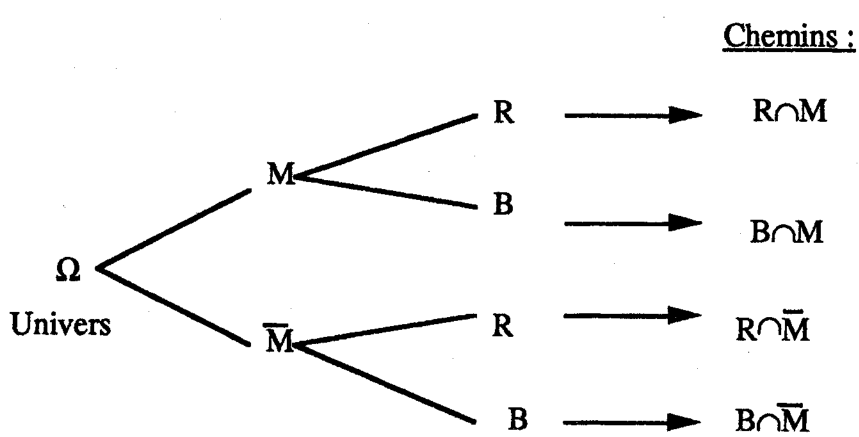
Je propose les représentations suivantes :

Il y a au moins deux possibilités d'analyser cette situation-problème à l'aide d'un arbre :

a)



b)



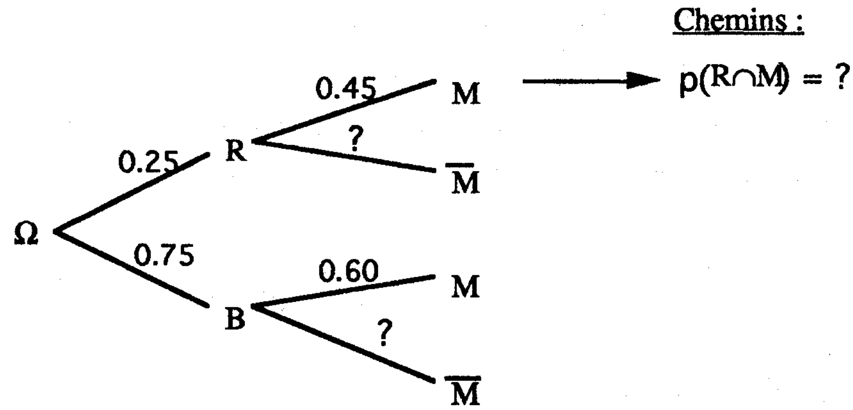
Convention à adopter :

Placer sur chaque branche de l'arbre la probabilité correspondante qui représentera le "poids" de la branche" sous la condition d'être passé par le nœud de son origine, la lecture s'effectuant de gauche à droite.

Une question :

Lequel de ces deux arbres peut-on remplir directement à partir de l'énoncé?

Evidemment, la bonne réponse attendue est l'arbre initial ou direct, c'est-à-dire la figure (**a**) ci-dessus :



Par définition (ou plutôt, par convention) :

Vu les conventions précédentes, les élèves ne peuvent pas faire autrement que proposer :

$p_R(M) = 0.45$, ou encore $p_R(M) = \frac{45}{100} = \frac{9}{20}$.

Sinon, quel sens donneraient-ils à R — M?

Je suggère, à cette occasion, d'amorcer l'institutionnalisation :

la probabilité de l'événement M, liée à la condition "*le carton contient une boîte rouge*" est $p_R(M) = 0.45$.

D'un point de vue dynamique, il est clair que c'est une probabilité de *transition* de l'événement R vers l'événement M, comme l'indique la figure R-M.

Pour conforter la précédente amorce de l'institutionnalisation, si le temps est suffisant, on pourra demander l'évaluation de chacune des probabilités conditionnelles suivantes :

$$p_B(M),\ p_B(\overline{M}) \text{ et } p_R(\overline{M}).$$

Ainsi, intuitivement, p_R et p_B prendraient déjà le statut d'une probabilité chez les élèves. Ces derniers ont quelque chance de prendre conscience du caractère relatif (à l'univers) d'une probabilité. Il ne reste plus que l'institutionnalisation véritable pour valider la "conjecture" intériorisée.

Vers la redécouverte de l'expression formelle d'une probabilité conditionnelle

a) Première question de mise en doute :

Evaluez les probabilités (qui figurent dans l'arbre direct) données dans l'énoncé sous forme de fraction irréductible.

La réponse attendue est: $p(R) = \frac{25}{100} = \frac{1}{4}$, et $p_R(M) = \frac{45}{100} = \frac{9}{20}$.

Sous l'hypothèse présente d'équiprobabilité (le tirage se faisant au hasard et les cartons étant identiques, au sens de "indiscernables à vue d'œil"), il est clair que 1/4 représente la fraction des cartons du magasin qui contiennent une boîte rouge (l'événement R).

b) <u>Deuxième question</u> :

Quelle est la fraction des cartons (R et M)?

La réponse attendue est : $\frac{9}{20}$ de $\frac{1}{4} = \frac{9}{20} \times \frac{1}{4} = \frac{9}{80}$.

Le premier membre de cette double égalité devrait être fourni presque spontanément par la classe elle-même comme je l'ai déjà remarqué lors du pré-test de l'année précédente (cf.[35] chap.III).

Il ne serait cependant pas étonnant que certains élèves proposent d'autres combinaisons d'opérations arithmétiques, comme par exemple : $\frac{25\%}{45\%}$.

c) <u>Troisième question</u> :

Quelle relation y a-t-il entre les trois probabilités $p(M \cap R)$, $p_R(M)$ et $p(R)$?

En cas d'hésitation prolongée, on précisera quelque peu la question.

Par exemple, on suggèrera :

Exprimez la probabilité $p(M \cap R)$ en fonction de $p_R(M)$ et $p(R)$

La réponse attendue est : $p(M \cap R) = \frac{9}{80} = p_R(M).p(R)$.

La notion de probabilité conditionnelle apparaît ainsi comme un outil pour le calcul de la probabilité composée $p(M \cap R)$. On profite de cette première mise au point pour fixer d'une manière anticipée quelques terminologies au sujet de l'utilisation d'un arbre des probabilités selon le schéma ci-dessous :

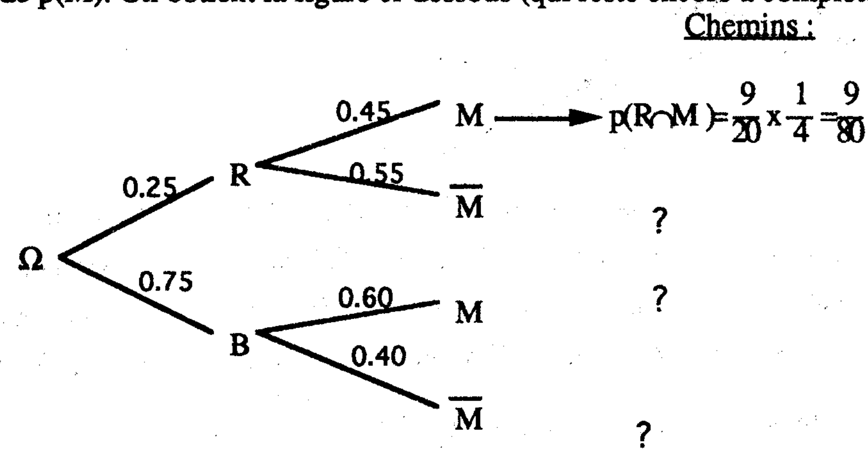
$$\underbrace{p(M \cap R)}_{\substack{\text{le poids du} \\ \text{chemin (R et M)}}} = \underbrace{\underbrace{p(R)}_{\substack{\text{le poids de la} \\ \text{branche} \\ \Omega \longrightarrow R}} \times \underbrace{p_R(M)}_{\substack{\text{le poids de la} \\ \text{branche} \\ R \longrightarrow M}}}_{\substack{\text{Produit des poids respectifs des branches} \\ \text{composantes du chemin (R et M)}}}$$

Précisons qu'il ne s'agit pas ici d'une institutionnalisation, mais d'un simple essai de traduction formelle mathématique d'une situation.

A supposer maintenant que seules les deux probabilités $p(M \cap R)$ et $p(R)$ soient connues, la probabilité $p_R(M)$ se calculerait alors par la formule : $p_R(M) = \frac{p(M \cap R)}{p(R)}$.

Vraisemblablement, on vient ainsi d'obtenir une nouvelle probabilité, p_R, à partir de la probabilité initiale p : c'est la *probabilité conditionnelle sachant R*. Remarquons également qu'une telle heuristique sur cette notion de probabilité conditionnelle permettrait d'éviter le classique recours préliminaire à un rapport de cardinaux, donc de retenir un éventuel *comportement cardinaliste*(2) qui risque de devenir réducteur, chez les élèves.

Afin de conforter ce nouveau concept, on propose ensuite de revenir à l'*arbre initial* des probabilités et de solliciter les élèves pour qu'ils complètent convenablement les diverses branches par leurs *poids* respectifs (en fait, c'est une situation à trous). On prépare ainsi le calcul de p(M). On obtient la figure ci-dessous (qui reste encore à compléter évidemment) :



Pour la question Q3 : détermination de p(M)

Vraissemblablement, on peut s'attendre que la classe trouve la solution de façon plus ou moins spontanée, par simple analyse de l'arbre des probabilités ci-dessus. Soit :

$$p(M) = p(M\cap R) + p(M \cap B) = \frac{9}{20} \times \frac{1}{4} + \frac{3}{5} \times \frac{3}{4} = 0.5175.$$

Ensuite, dans le souci de conforter la compréhension du mécanisme de calcul ainsi pratiqué, il est loisible de solliciter la classe à déterminer la probabilité $p(\overline{M})$ de l'événement complémentaire de M. On observera sans doute alors deux *modes différents* , c'est-à-dire :

- évidemment, par passage à l'événement contraire : $p(\overline{M}) = 1 - p(M)$;
- mais également, par calcul direct comme p(M), à savoir :

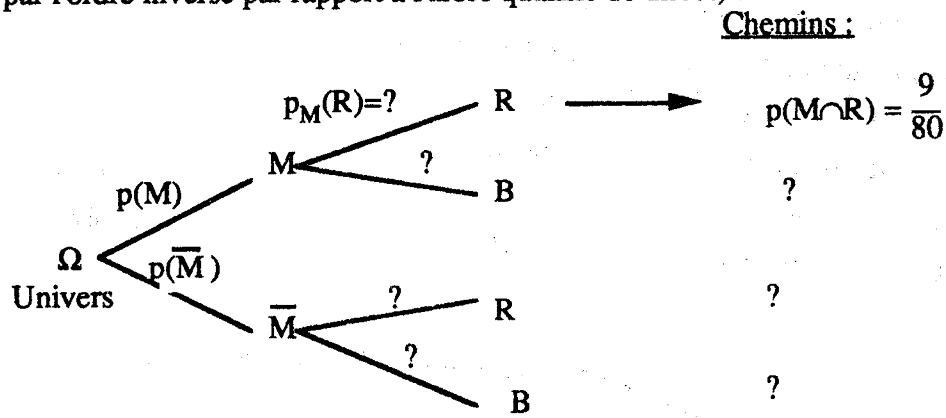
$$p(\overline{M}\cap R) + p(\overline{M}\cap B) = p(\overline{M}).$$

En ce qui concerne la question Q4

Il s'agit d'une question analogue à la question Q2. Seulement, les rôles joués ici par les événements R et M sont inversés. Ce qui me semble important ici, c'est la possibilité de renforcer la connaissance à faire acquérir aux élèves, c'est-à-dire une bonne perception du conditionnement, une attitude non cardinaliste sur une probabilité conditionnelle via une pratique de l'analyse arborescente. On invitera la classe, en cas de besoin , c'est-à-dire en cas

(2) Il s'agit du comportement qui consiste à croire qu'une probabilité conditionnelle $p_R(M)$ nécessite toujours la détermination du cardinal de $M\cap R$ et de celui de R (ce qui est possible ici d'ailleurs en fixant a priori l'effectif de Ω), comme nous l'avons déjà observé lors des tests de l'année 89/90 (cf.[35], chap. III]).

d'hésitation trop longue, à s'inspirer de ce qui vient d'être fait, c'est-à-dire utiliser une analyse arborescente pour résoudre le problème. Il s'agit alors de compléter l'*arbre inversé* (ou l'arbre obtenu par l'ordre inverse par rapport à l'arbre qualifié de direct) :



On remarquera que les deux événements $(M\cap R)$ et $(R\cap M)$ sont identiques, mais ceci peut poser problème aux élèves qui ont une conception chronologiste, donc non symétrique, pour une telle situation qui est d'ailleurs connotée d'une légère sémantique chronologique.

Cette fois, la classe sera nécessairement amenée à résoudre en x des équations du type

$$p(M\cap R) = x.p(M)\ ;\ \text{soit } x = p_M(R) = \frac{p(M\cap R)}{p(M)} = \ldots$$

Comme exercice d'entraînement, on pourra ensuite demander de calculer de deux manières les probabilités conditionnelles suivantes : $p_R(B)$, $p_{\overline{M}}(R)$, et $p_{\overline{M}}(B)$.

Petite synthèse en guise de conclusion pour cette phase d'introduction

Etant donné deux événements M et R dans un univers Ω fini , et une probabilité p définie sur cet univers, on a les relations du type :

$$p(M\cap R) = p(R\cap M) = p_R(M) \text{ x } p(R) = p_M(R) \text{ x } p(M);$$

Ce qui permet d'avoir :

$$p_R(M) = \frac{p(M\cap R)}{p(R)} = \frac{p_M(R) \text{ x } p(M)}{p(M\cap R)+p(R\cap\overline{M})}, \text{ si l'on connaît } p_M(R) \text{ et } p(M)\ ;$$

ou $$p_M(R) = \frac{p(M\cap R)}{p(M)} = \frac{p_R(M) \text{ x } p(R)}{p(M\cap R)+p(M\cap\overline{R})}, \text{ si l'on connaît } p_R(M) \text{ et } p(R).$$

C'est à partir de ce moment seulement qu'on se propose de donner le titre du chapitre en question, **Probabilité conditionnelle,** et de passer à la définition formelle classique, suivie de quelques propriétés.

Soit Ω un univers des possibles, et p une probabilité définie sur Ω. On procèdera ensuite ainsi: 1) Faire rappeler par les élèves la définition formelle d'une probabilité p.

2) Définition : Soit B un événement de probabilité non nulle.

De même qu'on a défini une probabilité p comme application

$$p : p(\Omega) \to [0, 1] :\ A \to p(A)$$

on appelle probabilité conditionnelle sachant B, l'application de $\mathcal{P}(\Omega)$ vers [0, 1] qui, à tout événement A, associe un nombre noté $p_B(A)$ (se lit p(A sachant B)) :
Cette application sera notée donc P_B :

$$p_B : \mathcal{P}(\Omega) \to [0, 1].$$
$$A \to p_B(A) = \frac{p(A\cap B)}{p(B)}.$$

Cette définition suppose que p_B est bien une probabilité.

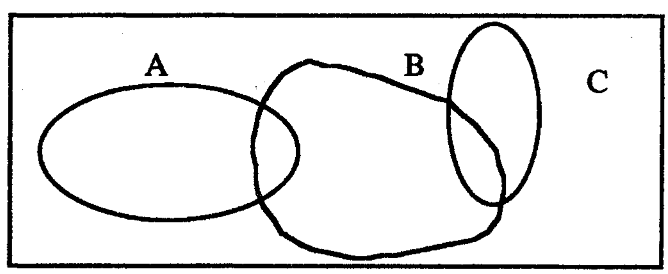
3) Faire vérifier que p_B est une probabilité effective. Admettre la distributivité de l'intersection par rapport à l'union (tout au plus, on fera une analogie avec la distributivité de la multiplication des réels par rapport à l'addition). On peut vérifier ce qui suit en exercice :

*** p_B est une probabilité sur Ω.**

En effet : 1°) $p_B(\Omega) = \frac{p(B\cap\Omega)}{p(B)} = \frac{p(B)}{p(B)} = 1.$

2°) Soit A et C sont deux événements incompatibles.

En s'aidant d'un support graphique , c'est-à-dire d'un diagramme de Venn ou de Carol [2], on admettra que $B\cap(A\cup C) = (B\cap A)\cup(B\cap C)$, avec $(B\cap A)$ et $(B\cap C)$ incompatibles.



Ω

$$P_B(A\cup C) = P(B\cap(A\cup C))/P(B) = P((B\cap A)\cup(B\cap C))/P(B)$$
$$= [P((B\cap A)+P(B\cap C)]/P(B) = P_B(A)+P_B(C).$$

Note :
On admettra alors que p_B satisfait les axiomes de probabilité, et donc qu'elle en possède toutes les propriétés.

En particulier : a) $p_B(A) = 1-p_B(A)$; $p_B(\emptyset) = 0$; $p_B(A\cup C) = p_B(A)+p_B(C)-p_B(A\cap C)$.

b) S'il y a hypothèse d'équiprobabilité pour la probabilité p, alors la probabilité conditionnelle se réduit à un simple rapport de cardinaux.

A titre de remarque à faire, on indiquera les autres notations qui sont aussi d'usage fréquent. Ainsi, il existe d'autres notations possibles qui ne seraient signalées aux élèves qu'à titre informatif : $p_B(A) = p(A/B)$, ou $p(A|B)$, en indiquant l'éventuelle conséquence infondée de croire que (A/B), ou (A|B) sont des événements(3) .

(3) Dans notre étude [35] sur les conceptions causaliste et chronologiste effectuée chez des étudiants de Deug, cette même année, quelques semaines avant le présent enseignement, nous avons pu remarqué l'existence d'un obstacle qui prendrait son origine dans la façon de noter la probabilité conditionnelle d'un événement.

Enfin, terminer par traiter un exemple simple, mais qui n'est pas d'un même modèle que celui présenté lors de l'étape introductive présentée précédemment :

Une urne contient 3 boules rouges et 2 boules noires. On tire successivement 2 boules de l'urne sans remise. 1°) Quelle est la probabilité de tirer 2 boules rouges?
2°) La première boule tirée est rouge. Quelle est la probabilité pour que la deuxième le soit aussi?

La première question est posée en vue de ne pas donner aux élèves l'impression de rompre complètement avec les connaissances "acquises" lors de la leçon qui précède celle-ci, c'est-à-dire le cas d'une probabilité discrète uniforme.

Soit R l'événement "*tirer 2 boules rouges*". R est l'ensemble des combinaisons à 2 éléments de l'ensemble des 3 boules rouges.

$$\mathrm{Card}R = C_3^2 \quad \text{et } P(R) = \frac{C_3^2}{C_5^2} = \frac{3}{10}.$$

Pour la deuxième question, on insistera surtout sur le rôle de l'*information supplémentaire*, ou du conditionnement de l'événement inféré par la succession de tirages effectués sans remise. On laisse les élèves faire des propositions sur la possibilié d'évaluer la probabilité conditionnelle en question de deux manières, et ce sans utilisation systématique de l'analyse arborescente. Avec les notations d'événements telles que :

R1 : 'la 1ère boule tirée est rouge " ; R2 : "la 2ème boule tirée est rouge",
il s'agit de calculer $p_{R1}(R2)$. En fait, il y a deux méthodes ici :

Méthode directe :
A priori, on a $p_{R1}(R2) = 1/2$. A cette occasion, il est loisible de redemander la probabilité de l'événement $R1 \cap R2 = R$, pour faire constater la différence avec $p_{R1}(R2)$, ou $p_{R2}(R1)$.

Deuxième méthode :
Par définition d'une probabilité conditionnelle, $p_{R1}(R2) = \frac{p(R1 \cap R2)}{p(R1)} = \frac{p(R)}{p(R1)}$.
Or $p(R) = 3/10$ et $p(R1) = 3/5$. D'où $p_{R1}(R2) = \frac{3/10}{3/5} = \frac{1}{2}$.

On termine ensuite par le renforcement suivant pour le calcul de p(R) :
$p(R) = p(R1 \cap R2) = p_{R2}(R1) \times p(R2) = p_{R1}(R2) \times p(R1)$ (de préférence ici...)"

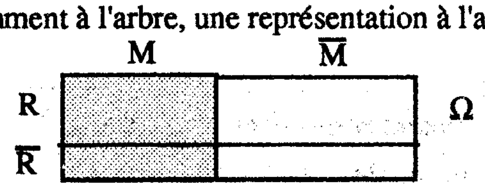
VI Bilan qualitatif d'une expérimentation

L'expérimentation faite auprès des élèves d'une classe de terminale D[1] du lycée de Bréquigny à Rennes en 1991 a permis de dégager les points suivants (cf.[35] chap.V, pour le détail) :

1) l'identification d'un pourcentage à une probabilité n'est pas spontanée. Elle nécessite un

[1] Cette classe est tenue sous la responsabilité de Monsieur A.SIMON, un collègue de recherche, membre de l'équipe didactique de l'I.R.MA.R. Je lui remercie d'avoir accepté de jouer le rôle d'acteur didactique, sans degré de liberté. Ce qui m'a permis de mieux observer les réactions des élèves.

apprentissage supplémentaire chez les élèves, même s'ils ont acquis une certaine notion de probabilité. Cette résistance peut tenir au fait que le pourcentage est connoté par les contextes concrets où il est la mesure d'une partie d'un tout ;
2) la conception *cardinaliste* apparaît comme un obstacle à l'acquisition de la notion de probabilité conditionnelle ;
3)l'utilisation d'un arbre des probabilités comme signifiant figural offrirait une possibilité d'intégrer *en acte* l'apprentissage de la formule de BAYES dès le niveau de la terminale, sans demander un coût cognitif énorme ;
4) aucun indice de conception causaliste de cette notion ne s'est manifesté ;
5) par rapport à d'autres méthodes d'introduction de la notion de probabilité conditionnelle (motivation par fréquences conditionnelles...), celle consistant à partir du caractère intuitif de la formule de probabilités composées n'apparaît pas absurde ; mais il reste quand même à trouver le bon contexte.
6) enfin, le plus important est que l'analyse d'un texte par dichotomie arborescente semble favoriser la conception chronologiste de cette notion, d'autant plus que la conjonction *et* peut induire une chronologie entre les événements connectés. C'est là un défaut de cette procédure par dichotomie arborescente. Elle est vraisemblablement associée à la conception chronologiste de la probabilité conditionnelle, une conception dont le renforcement devient un obstacle didactique. Pour dépasser cet obstacle, il nous semble maintenant utile d'alterner avec un autre mode de représentation de la situation. Pour être clair, il serait possible d'indiquer, concurremment à l'arbre, une représentation à l'aide d'un "tableau à double entrée",



ou tout simplement à l'aide d'une "patate" comme dans la pratique traditionnelle. On verra plus tard les difficultés levées par cette représentation : la probabilité conditionnelle est lue comme une probabilité conjointe. Toutefois, tous ces points demeurent hypothétiques. Seule l'analyse des productions d'élèves sur des problèmes suffisamment variés peut éventuellement conforter ces premières impressions.

VII Evaluation immédiate : le lendemain de la séance du cours

VII-A Evaluation 1 :

Enoncé : Fabrication des boulons

Une usine dispose de deux machines M1 et M2 fabriquant des boulons.
La machine M1 fabrique 40% de la totalité des boulons ; 5% des boulons fabriqués par cette machine sont défectueux ; 1% des boulons fabriqués par M2 est défectueux.
On tire un boulon au hasard. Tous les boulons ont la même probabilité d'être tirés.
L'examen du boulon tiré montre qu'il est défectueux. Quelle est la probabilité qu'il ait été fabriqué par la machine M1.

Brève analyse liminaire de l'énoncé

Il s'agit d'une situation-problème isomorphe à celle traitée collectivement pendant la séance d'introduction du chapitre sur la notion de probabilité conditionnelle. La seule différence est que les questions préliminaires sont absentes cette fois. Les élèves ont la possibilité de consulter leurs notes manuscrites. Pour justifier un tel choix de situation, mon argument est simple. Je me place dans une situation classique d'enseignement où le premier test doit comprendre au moins un exercice qui se prête à une application directe du contenu du cours suivi. Justement ici, comme je viens de le dire, il est à remarquer que seul l'habillage a changé par rapport à l'énoncé du problème d'introduction de la leçon. Il est le plus "familièr" pour que l'éventuel équilibre existant chez les élèves ne soit pas compromis.
Vu la nature du problème traité pendant la séquence d'enseignement qui a précédé ce test, je pense qu'une indication sur la décomposition sus-mentionnée n'est plus nécessaire, de même pour la question préliminaire. Concernant la prévision des démarches déployées par les élèves, mon attente repose sur une utilisation massive de l'arbre des probabilités aboutissant à une réponse pertinente. Toutefois, il n'est pas impossible de retrouver l'attitude *cardinaliste*, à savoir la considération d'un échantillon de 100 boulons par exemple, comme je l'ai déjà signalé au début de la première séquence d'enseignement de la notion de probabilité conditionnelle.

VII-B Evaluation 2

Enoncé : Problème d'urnes

Deux urnes U1 et U2 contiennent respectivement 2 boules blanches et 3 boules noires, et 4 boules blanches et 1 boule noire. On tire au hasard une des urnes, puis on tire de cette urne une boule : on obtient une boule blanche. quelle est la probabilité que cette boule provienne de l'urne U1?

Brève analyse liminaire de l'énoncé

Il s'agit d'un cas particulier du modèle d'urnes du problème abusivement dénommé des "probabilités de causes ou d'hypothèses". Deux "causes" (les deux urnes), avec deux "effets" possibles (couleur blanche ou noire de la boule tirée) pour chaque "cause" sont présentées ici. Comme démarches de réponses attendues, on peut prévoir aussi bien l'utilisation que la non-utilisation d'un arbre des probabilités. Quelle est la procédure qui favorise le plus la réussite? Le premier cas montrerait la capacité de l'élève à faire une transposition de la méthode arborescente qui vient d'être *institutionnalisée* dans une situation-problème où les probabilités sont exprimées en pourcentages. Le second cas serait dû à la résistance présentée par le calcul basé sur le rapport "*nombre des cas favorables sur nombre des cas possibles*", c'est-à-dire dû à un *comportement cardinaliste*, surtout pour la détermination de la probabilité de l'événement conditionnant. En effet, les deux urnes étant équiprobables, on peut considérer comme univers des possibles l'ensemble des 10 boules (le mélange des contenus des deux urnes), celles-ci étant indexées par l'urne qui les contient, soit $W = \{b_{11}, b_{21}, n_{11}, n_{21}, n_{31}, b_{12}, b_{22}, b_{32}, b_{42}, n_{12}\}$. Ainsi, dans ce cas la probabilité de l'événement "obtenir une boule blanche" est égale à 6/10. Mais, il serait possible aussi de rencontrer une méthode qui consiste à faire le rapport 2/6 =1/3 , où $2=\mathrm{card}(\{b_{11}, b_{21}\})$ et $6=\mathrm{card}(\{b_{11}, b_{21}, b_{12}, b_{22}, b_{32}, b_{42}\})$.

En résumé, je m'attends à deux démarches :

- une démarche qui traduit un comportement *cardinaliste* : elle consiste ici, nous le repétons, à évaluer la probabilité conditionnelle comme rapport des cardinaux :

$p_B(U_1) = \dfrac{\mathrm{card}(U1 \cap B)}{\mathrm{card}B}$, formule qui n'est valable que dansle cas d' hypothèse d'équiprobabilité des urnes U1 et U2.

- une démarche plutôt *fondamentaliste*, c'est-à-dire qui mobilise la définition donnée d'une probabilité conditionnelle et la formule des probabilités composées :

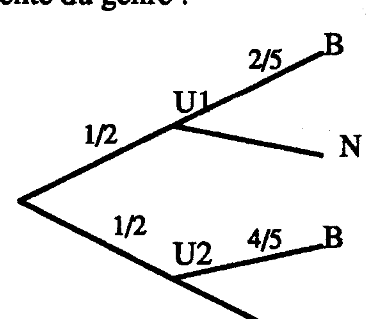
$$p_B(U_1) = \frac{p(U_1 \cap B)}{p(B)} = \frac{p_{U_1}(B).p(U_1)}{p(B)}, \text{ avec } p(B) = p(B \cap U_1)+p(B \cap U_2) ;$$

or $p(B) = p_{U_1}(B).p(U_1)+p_{U_2}(B).p(U_2) = \frac{1}{2} \times \frac{2}{5} + \frac{1}{2} \times \frac{4}{5} = \frac{3}{5}$;

D'où la probabilité que la boule blanche tirée provienne de l'urne U1 est :

$$p_B(U_1) = \frac{\frac{1}{2} \times \frac{2}{5}}{\frac{3}{5}} = \frac{1}{3} \approx 0.33.$$

Enfin, chacune de ces deux procédures de résolution pourrait également s'accompagner d'une analyse arborescente du genre :



$\longrightarrow p(B \cap U1) = \frac{1}{2} \times \frac{2}{5} = \frac{1}{5}$

$\longrightarrow p(B \cap U2) = \frac{1}{2} \times \frac{4}{5} = \frac{2}{5}$

VIII Analyse des démarches effectives de réponses des élèves

VIII-A Sommaire

Faute d'une autre classe de la même série, ces deux exercices, *Evaluations 1 et 2*, sont testés également auprès des élèves de la terminale A1 (tenue sous la responsabilité de Madame A.LARHER, collègue de recherche) du Lycée Ile-de-France qui venaient de recevoir un enseignement sur la notion de probabilité. Mais la situation didactique menée dans cette classe ne mobilisait pas la variable arborescence d'une manière institutionnalisée comme dans la classe de terminale D du lycée de Bréquigny où se passait l'expérimentation véritable. Aussi, pour indiquer brièvement une comparaison entre les performances globales des deux classes, je donne ci-dessous un tableau de réussites-échecs relatifs à ces deux situations:

Classes	FABRICATION DES BOULONS		MODELES D'URNES		Effectifs
	Réussite	Echec	Réussite	Echec	
TA	**5**	38	**13**	30	**43**
TD	**20**	12	**9**	23	**32**

Il serait tentant de conclure, hâtivement, que les rapports de réussites à ces deux problèmes sont quasiment inversés dans les deux classes. Il paraît naturel de se poser la question sur les éventuels liens de ce phénomène avec les variables et hypothèses didactiques fixées a priori, et avec d'autres conceptions sous-jacentes à certaines procédures déployées par les élèves à travers leurs réponses.

Pour le reste des évaluations prévues, l'analyse sera faite un peu plus loin, car cela ne concerne que la TD.

VIII-B Analyse statistique de quelques croisements de réussites-échecs

VIII-B.1 : Classe TD

Tableau de Réussites-Echecs aux deux problèmes du premier test

		Fabrication des boulons		
		Réussite	Echec	
Problème d'urnes	Réussite	12	**0**	37,5% 12
	Echec	**8**	12	62,5% 20
		20 62,5%	12 37,5%	32

$\chi^2_{\text{Mac-Némar corrigé}} \approx 6.12$ ($\chi^2_{\text{lu}} = 3.84$ à 5% à 1 degré de liberté) :

Il en résulte que ce test de χ^2 est significatif au seuil de 5%. Il y aurait donc *dépendance entre les résultats à ces deux problèmes*. Vraisemblablement, le "problème d'urnes" apparaît plus difficile que l'autre problème. Ce résultat se trouve conforme à notre attente. Le problème demeure maintenant de faire investir cet outil ou cette méthode, c'est-à-dire l'arborescence, pour attaquer des problèmes relevant de la notion de probabilité conditionnelle où l'information supplémentaire n'apparaît pas aussi évidente.

Croisements de "l'arborescence" avec les deux problèmes

		Fabrication des boulons		
		Réussite	Echec	
Arborescence	Présence	19	11	93,7% 30
	Absence	1	1	6,3% 2
		20 62,5%	12 37,5%	32

$\chi^2_{\text{Mac-Némar corrigé}} \approx 6.7$

		Problème d'urnes		
		Réussite	Echec	
Arborescence	Présence	11	14	78,1% 25
	Absence	1	6	21, 9% 7
		12 38%	20 62, 3%	32

$\chi^2_{\text{Mac-Némar corrigé}} \approx 9.6$

Ces deux tableaux nous indiquent une légère diminution de la pratique de l'analyse arborescente pour le problème d'urnes où ne figurent plus des pourcentages. Ce phénomène est tout à fait normal dans un début d'apprentissage. Il demeure encore une forte prégnance presque naïve du cours que les élèves viennent de suivre. La preuve est que certains élèves tentent de ramener les nombres figurant dans l'énoncé du "problème d'urnes" à des pourcentages (cf. les exemples de productions écrites relevées dans [35] chap.V).

Analyse implicative et cohésitive de R.GRAS et A.LARHER

Remarque méthodologique pour le traitement d'un croisement de réussite-échec
Il est clair que les trois croisements analysés ci-dessus selon le test traditionnel de χ^2 font déjà ressortir des hypothèses d'implications; cependant, de telles hypothèses demeurent essentiellement séquentielles, car on est obligé de procéder séparément à toutes les combinaisons possibles des variables prises deux à deux. La méthode d'analyse implicative présente ici une différence par rapport à un test de χ^2 pour l'étude des implications; outre son aspect multidimensionnel, elle peut trancher sur la transitivité, à cause de la relation de préordre partiel associée (cf. [19], [20] ou [21]). Il convient donc de procéder à une analyse implicative et cohésitive (voir la description intuitive et succincte présentée dans l'annexe, pp.29-31).

Rappelons les 6 variables de réussite-échec concernées ici :

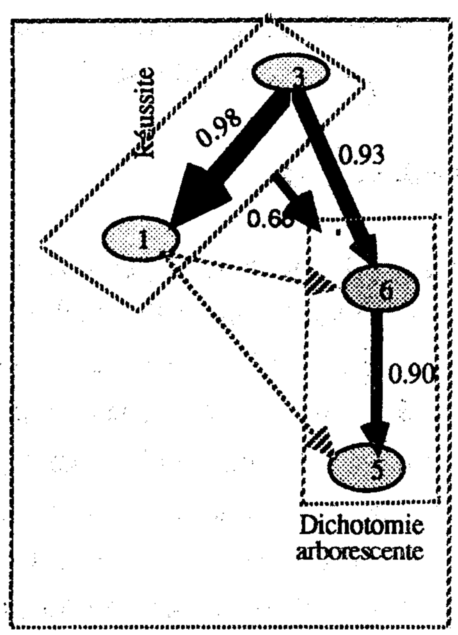
Variable **1** : la réussite au problème *fabrication des boulons*;
Variable **2** : l'échec au problème *fabrication des boulons;*
Variable **3** : la réussite au problème *modèle d'urnes*;
Variable **4** : l'échec au problème *modèle d'urnes*;
Variable **5** : utilisation de l'arborescente au problème *fabrication des boulons*;
Variable **6** : utilisation d'une analyse dichotomique arborescente au problème *modèle d'urnes*;

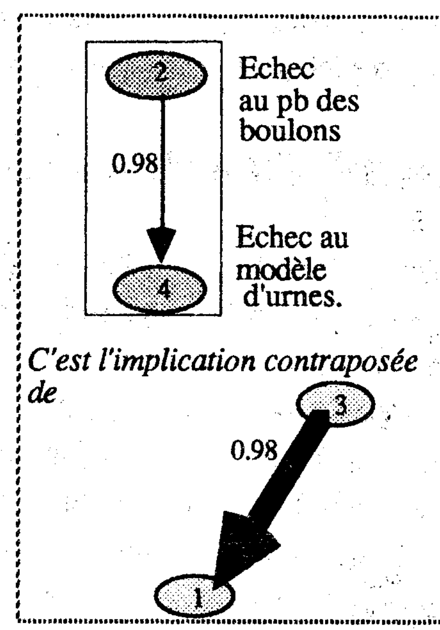
Tableau des résultats numériques relatifs à ces six variables :

Intensités d'implication et *cohésions des couples orientés*

sens de l'implication ↗		1	2	3	4	5	6
intensité d'implication	1	1			0.401	0.568	0.617
cohésion		*1*			*0*	*0.163*	*0.280*
"	2	0.0004	1		0.983	0.418	0.359
		0	*1*		*0.992*	*0*	*0*
"	3	0.983	0.04	1	0.000	0.796	0.925
		0.992	*0*	*1*	*0*	*0.684*	*0.922*
"	4				1	0.274	0.149
					1	*0*	*0*
"	5					1	
						1	
"	6					0.898	1
						0.881	*1*

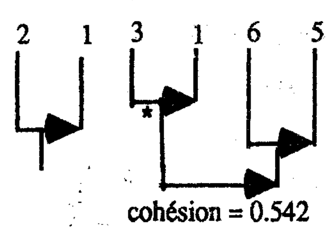
Voici le graphe implicatif[1] obtenu au seuil de 0,90 d'intensité d'implication :





cohésion = 0.542, valeur théoriquement étonnamment grande pour une classe implicative de 4 variables.

Représentation linéaire obtenue par l'arbre de classification implicative et cohésitive:



Remarquons que cette classification hiérarchique implicative et cohésitive est exceptionnellement en parfaite harmonie avec le graphe implicatif correspondant.

(1) Le traitement des données est fait sous le logiciel C.H.I.C. élaboré dans le laboratoire de didactique de l'I.R.MA.R.

En fait, compte tenu de l'équivalence logique (a $\Rightarrow$b) $\Leftrightarrow$ (non b $\Rightarrow$ non a), pour deux variables antagonistes, il suffit de prendre une d'elles (par exemple celle qui est, a priori ou non, en corrélation positive avec certaines du reste des variables) dans le tableau des données. L'autre implication duale de celle obtenue directement par la lecture du graphe implicatif s'obtient par simple contraposition. En retenant la variable *réussite,* par exemple, on tient en compte implicitement l'*échec* correspondant. Ce qui justifie mon propos ci-dessus. En ne retenant que les variables 1 et 3 , par exemple, on obtient le graphe implicatif présenté à gauche ci-dessus. L'implication entre classes $\{3, 1\} \Rightarrow \{6, 5\}$, sachant que la valeur de la cohésion implicative de la classe union $\{3, 1, 6, 5\}$ est très significative[1] , donne avec pertinence un résumé de la relation existant entre la procédure arborescente et la réussite aux deux problèmes bayésiens. Le graphe implicatif complète l'information ainsi recueillie : la forte implication $\{3\} \Rightarrow \{6\}$ apparaît quelque peu surprenante, vu le type de la situation exploitée lors de la séquence d'enseignement en TD. Cela corrobore l'implication entre classes ci-dessus.

Interprétation

Le graphe implicatif des réussites, hormis l'ordre de complexité des deux problèmes pour ces élèves de TD, nous revèle que même pour la situation qui, a priori, n'a rien à voir avec la situation d'enseignement, l'analyse arborescente demeure comme une condition nécessaire à la réussite. Soit de façon contraposée, ne pas pratiquer l'analyse arborescente suffit à conduire à l'échec (partiellement pour le problème d'urnes).

VIII-B.2 : Classe témoin TA

Nous nous contentons ici de faire une analyse implicative relative aux 4 variables suivantes :

1 : la réussite au problème *fabrication des boulons ;*

2 : utilisation d'un *tableau à double entrée* comme support visuel des événements à ce dernier

3 : la réussite au problème *modèle d'urnes ,*

4 : utilisation d'un *tableau à double entrée* comme support visuel des événements à ce dernier.

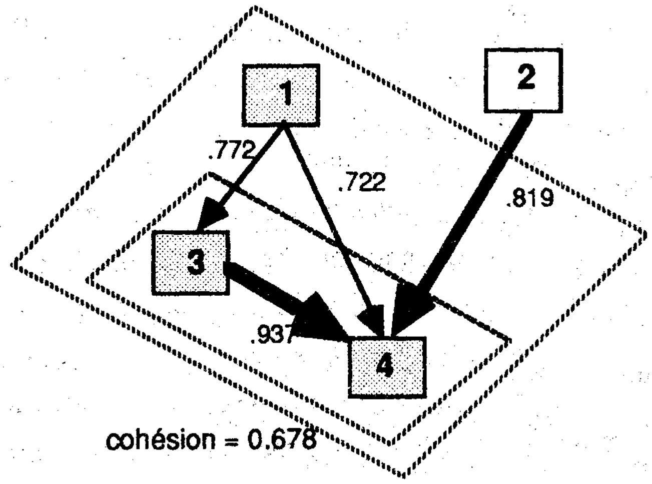
Signalons que l'utilisation d'un support visuel autre qu'un tableau à double entrée est quasiment inexistante chez les élèves de TA lors de ce test (soit un élève sur 43).

---

1 Théoriquement (cf.[36] chap.I), les seuils de confiance de cohésion implicative d'une classe se répartissent ainsi :

- pour une classe à deux variables, la cohésion dépasse 0.95 avec une probabilité égale à 0.05 ;
- pour une classe à trois variables, la cohésion dépasse 0.62 avec une probabilité égale à 0.05 ;
- pour une classe à 4 ou 5 variables, la cohésion dépasse 0.400 avec une probabilité inférieure à 0.03.

Graphe implicatif obtenu au seuil de 0.72 d'intensité d'implication :



Il est à remarquer que la valeur 0.72 n'est pas très significative pour une intensité d'implication entre deux variables. Néanmoins, la forte valeur 0.678 de cohésion de la classe $\{1, 3, 4\}$, où $\{1\} \Rightarrow \{3, 4\}$, rend légitime la considération de ce graphe (voir la note, p.23).

Interprétation

- d'un point de vue ordre de complexité, contrairement à la classe TD, chez ces élèves de TA le problème *fabrication des boulons (évaluation 1)* serait nettement plus difficile que le problème *de modèle d'urnes ;*
- d'un point de vue procédure de résolution, l'utilisation d'un tableau à double entrée comme support visuel des événements favoriserait d'une façon très nette la réussite à un problème de type *modèle d'urnes* , mais d'une façon moins significative que pour le problème du genre *fabrication des boulons.*

VIII-B.3 : Conclusion

La dissemblance des résultats obtenus, à la suite de la mobilisation de deux "outils"différents, par rapport aux réussites et échecs à ces deux premiers problèmes pour les deux classes TA et TD, qui sont tenues par deux enseignants, distincts certes, nous autorise à conjecturer que l'utilisation d'un arbre des probabilités et celle d'un "tableau à double entrée" (ou un diagramme de Venn [2]) sont intégrables comme variables didactiques quand s'il s'agit d'introduire la notion de probabilité conditionnelle.

L'ingéniérie didactique expérimentée ici, basée sur la notion de pourcentages et sur l'utilisation d'un arbre des probabilité, offrirait une autre voie non nécessairement dépourvue d'avantages par rapport à d'autres modes, tel le classique modèle d'urnes, pour introduire le concept de probabilité conditionnelle et la formule de Bayes.

En revanche, il serait très intéressant également de "creuser" plus profondément ce qui sous-endrait les échecs.

VIII-C Quelques éléments d'explication sur les démarches non pertinentes

Je me contenterai ici à l'évaluation 1 (Pour la consultation de quelques copies d'élèves, voir [36] pp.239-260). Grosso modo les erreurs les plus marquantes portent sur deux points :

VIII-C.1-les difficultés procédurales

En plus d'adopter une démarche métamathématique (par opposition à une démarche formelle, c'est-à-dire : l'univers des possibles reste non ou mal défini, les notations d'événements sont implicites), certains élèves tentent de s'aider d'un support graphique pour mieux appréhender la situation-problème. Par exemple, à propos du "problème de boulons", examinons les deux réponses proposées par deux élèves de terminale A1 ci-dessous :

Réponse 1 :

	M_1	M_2
bons boulons	..	
boulons défectueux	5/100	1/100
	4/10	6/10

A = "être des boulons de M_1" $p(A) = \frac{4}{10}$

$\bar{A}$ = " ——— M_2" $p(\bar{A}) = \frac{6}{10}$

B = "être 1 bon boulon" $\frac{94}{100} = p(B)$

$\bar{B}$ = " — 1 boulon defecheux" $p(\bar{B}) = \frac{6}{100}$

$$P_{\bar{B}}(A) = \frac{p(\bar{B} \cap A)}{p(\bar{B})}$$

Une telle réponse témoigne un besoin de se représenter graphiquement la situation, en l'occurrence par un tableau à double entrée. Mais comme l'attestent les deux cases non remplies (en réalité cela correspond à une trace de gomme), cet élève ne peut placer convenablement tous les nombres qui apparaissent dans l'énoncé. Le fait de placer 5/100 (i.e.5%), par exemple, dans la case intersection de a la ligne "boulons défectueux " avec la colonne "M1" laisse penser qu'il y a la confusion entre probabilité conditionnelle et probabilité conjointe. Ce mode de représentation figurale lui apparaît non adapté au conditionnement d'événément .

Réponse 2 :

M_1 fabrique 0,4 de la totalité des boulons

60+10=100 %

parmi eux 0,5 sont défectueux

M_2 fabrique que 0,6 de la totalité des boulons

= 60%

0,1 sont défectueux

	M_1	M_2
D	5%	1%
$\bar{D}$	95%	99%
	0,4	0,6

100%

Défectueux = D

$$p(M_1 \cap D) + p(M_2 \cap D) = p(D)$$

$$\cancel{P_{M_1}(D) = \frac{P(M_1 \cap D)}{P\,M_1} = \frac{}{0,4}}$$

boulon défectueux = proba conditionnelle

$$P_D(M_1) = \frac{P(D \cap M_1)}{P(D)} = \frac{P(D \cap M_1)}{P(M_1 \cap D) + P(M_2 \cap D)}$$

On constate ici que, malgré la pertinence de la formule des probabilités totales et de la probabilité conditionnelle à évaluer, l'élève auteur de la "Réponse 2" se trouve bloqué. Cela provient-il tout simplement d'une mauvaise identification des informations?

On peut formuler diverses hypothèses pour expliquer ce type de blocage. En tout cas, la représentation ainsi choisie, en l'occurrence un tableau à double entrée, s'avère *non-congruente*(*) à cette situation. G.VERGNAUD[38] avait déjà montré que "la fonction des signifiants langagiers et symboliques est de favoriser la conceptualisation (identification et objectivation) et de permettre la communication et le débat ; ils jouent en outre un rôle de régulation de l'action dans la résolution de problème. La diversité des signifiants, des signifiés et des référents pose le problème de la congruence sémantique : tous les signifiants ne se valent pas, certains reflètent mieux que d'autres les propriétés du signifié".

L'un des "vecteurs orientant la complexité" retenus par R.GRAS(cf.[19] pp.4-5), partie: *Point d'appui de la nouvelle taxonomie)* pour construire "la taxonomie" spécifique mathématique, est la "séparabilité significative du signifiant et du signifié". Le problème qui se pose est : comment mesurer cet écart signifiant-signifié?

Rappelons que d'une manière générale, on parle concurremment de plusieurs types de registres en mathématiques, à savoir le registre (au sens de DUVAL [11], [12]) de notations symboliques (ou formalisme), le registre langagier mathématique, le registre langagier naturel, le registre figural. A.MESQUITA a montré dans sa thèse[28] que dans une tâche mathématique, en géométrie en l'occurrence, les informations issues de ces registres peuvent être soit concordantes, cas où il y a une congruence sémantique entre les registres mobilisés (les objets sont alors identifiables par simple appréhension), soit non concordantes, cas de *non-congruence sémantique* entre les registres en question. En fait, on distingue deux types de congruence : une, *inter-registre*, entre registres différents ; une deuxième, interne à un registre, *intra-registre*. Il importe donc de considérer l'articulation entre différents registres pour hiérarchiser, ou pour expliquer, la complexité d'une tâche mathématique. Pour une étude plus détaillée de ce jeu d'articulation inter ou intra-registre, nous renvoyons à l'article de DUVAL[26] et à la thèse de A.MESQUITA[28].

Néanmoins, pour ma part, il semble que ce conflit, disconcordance inter-registres, apparaît dans la résolution d'une situation-problème spécifique de la notion de probabilité conditionnelle. Par exemple, dans le cas d'un tableau à double entrée tel que les élèves le construisent ici, la forme d'appréhension séquentielle (qui est en accord avec le conditionnement) s'oppose à l'appréhension perceptive de l'intersection (ou d'événements conjoints) directement suggérée par un tel tableau : ce qui explique la non-congruence sémantique entre ce mode de représentation figurale et un problème de probabilité

(*) Ce terme de congruence (ou de non-congruence) sémantique est utilisé ici dans un sens tout à fait analogue à ce qui est dit par DUVAL(1988), pour la première fois dans la théorie didactique à notre connaissance, et poursuivi par A.MESQUITA (Thèse, 1988, chapitre 1er) à propos de l'apprentissage de la démonstration mathématique chez des jeunes élèves.

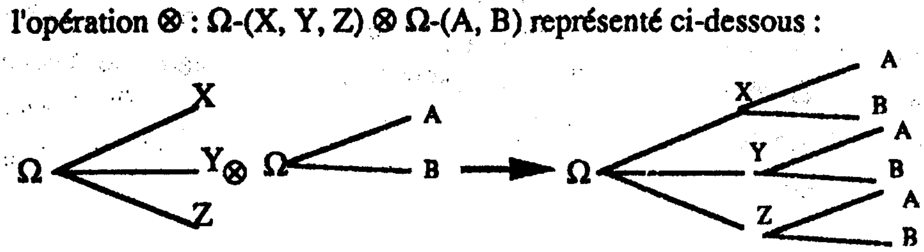
conditionnelle. En effet, le tableau à double entrée est communément lu de façon symétrique, alors que le conditionnement est non symétrique. Ce qui explique aussi en partie l'échec des élèves pour ce "problème de fabrication des boulons" comme l'attestent leurs productions écrites. Le tableau à double entrée apparaît ici en quelque sorte comme un piège de l'*intuition*. Alors que, contrairement à cela, comme nous venons de le voir dans le paragraphe précédent, l'exploration d'un arbre des probabilités, qui suggère une perception séquentielle d'événements, joue un rôle à la fois *descriptif* et *heuristique* pour ce qui concerne la résolution du présent problème : une représentation arborescente apparaît alors comme sémantiquement congruente à la notion de probabilité conditionnelle, mais risque d'ériger un obstacle dans les problèmes bayesiens. En conséquence, une prudente réserve didactique s'impose.

Cette notion de congruence sémantique, semble expliquer en partie une conclusion de S.Maury(thèse [27], 1986) qui affirme que la difficulté de résolution d'un problème de probabilité élémentaire *dépend du contexte en jeu, du vocabulaire et de la façon de formuler la question*. En effet, si le contexte qui sous-tend le problème est par exemple en accord avec le registre langagier naturel (i.e. de la vie quotidienne, conforme avec la rationalité quotidienne), son rôle heuristique sera différent de celui d'un contexte formel.

D'une manière plus rigoureuse, la théorie des représentations calculables exige l'existence des homomorphismes [40] du système représenté dans le système représentant. Ce sont de tels homomorphismes qui diminuent le coût cognitif lors d'un processus d'apprentissage.

Or, pour ce qui concerne le concept de probabilité conditionnelle, on peut démontrer que parmi les deux signifiants candidats, l'arborescence et le tableau à double entrée, le premier permet de construire un homomorphisme signifié - signifiant. Car l'arborescence est orientable, donc compatible avec une dissymétrie qui caractérise justement le concept en question ici; au contraire, avec un "tableau", c'est plutôt la symétrie ou la commutativité (chaque case représentant une intersection). Concrètement, désignons par Partition(Ω) l'ensemble de toutes les partitions de l'univers Ω, et par Arbre(Ω) celui des arbres d'origine Ω, construits à partir des événements de Ω. Définissons les lois de composition suivantes :

- dans Partition(Ω) : par exemple, pour deux éléments (X, Y, Z) et (A, B) de Partition(Ω), définissons $(X, Y, Z)\oplus(A, B) = (X\cap A, X\cap B, Y\cap A, Y\cap B, Z\cap A, Z\cap B)$ où l'intersection est supposée ordonnée;
- dans Arbre(Ω) : par exemple, pour deux arbres Ω-(X, Y, Z) et Ω-(A, B), définissons l'opération $\otimes$: Ω-(X, Y, Z) $\otimes$ Ω-(A, B) représenté ci-dessous :



Alors l'application f définie de Partition(Ω) vers Arbre(Ω) par f(A, B) (notation simplifiée de f((A, B))) est l'arbre Ω-(A, B), avec la contrainte suivante : f($X\cap A$, $X\cap B$, $Y\cap A$, $Y\cap B$, $Z\cap A$,

$Z \cap B$) est justement l'arbre ci-dessus, est bien un homomorphisme :

pour toutes partitions $a=(A_1,\ldots, A_k)$, $b=(B_1,\ldots, B_r)$ de Ω, on a $f(a \oplus b) = f(a) \otimes f(b)$.

VIII-C.2- les obstacles conceptuels

Ils se situent essentiellement à trois niveaux :

- l'évaluation de la probabilité de l'événement conditionnant ;
- la confusion entre probabilité d'une intersection et probabilité conditionnelle ; plus précisément, une telle confusion s'effectue dans les deux sens à savoir : prendre la probabilité conditionnelle à la place de la probabilité composée, et inversement, considérer la probabilité conjointe au lieu de la probabilité conditionnelle ;
- l'obstacle de renversement de situation qui consiste à inverser le rôle de l'événement conditionné et celui de l'événement conditionnant, en l'occurrence considérer $p_{M1}(D)$ au lieu de $p_D(M1)$. Ce type d'obstacle peut s'expliquer par une éventuelle conception chronologiste ou causaliste de la notion de probabilité conditionnelle. En effet, on peut penser que, dans l'écriture $p_A(B)$, l'événement B ne jouerait plus chez l'élève (causaliste) le rôle d'une simple variable, et que A serait un *générateur de conséquences*. Mais, il me semble relever d'un obstacle psycho-cognitif plus général, celui de la *réversibilité*. Je cite R.GRAS ([19] pp.15-16) à ce sujet :

"Le sujet ne se connaît psychologiquement et biologiquement (adéquation et de l'esprit et du corps) qu'en s'adaptant, en assimilant l'objet et en s'y accomodant comme le fait l'œil lors des visées à des distances variées...Mais ces régulations ne se distinguent véritablement des perceptions que dans la mesure où elles sont susceptibles de composition et de réversibilité. C'est en cela que l'acte intelligent conscient se distingue de l'acte réflexe. Autrement dit, c'est dans l'aptitude à opérer selon la fonction réciproque d'une action directe, annulant son effet, que se revèlent les facultés accommodatrices, signes de l'intégration entière aux schèmes sensori-moteurs du sujet, lesquels constituent un ensemble d'actions intériorisées...L'homomorphisme de transposition, qui permet la substitution d'une action par un autre, doit être réversible pour être significatif de l'accès à la représentation en tant que telle". Ainsi, c'est cette carence de réversibilité relativement au conditionnement qui fournit au mieux un élément d'explication de la résistance des conceptions chronologiste ou causaliste au concept formel de probabilité conditionnelle.

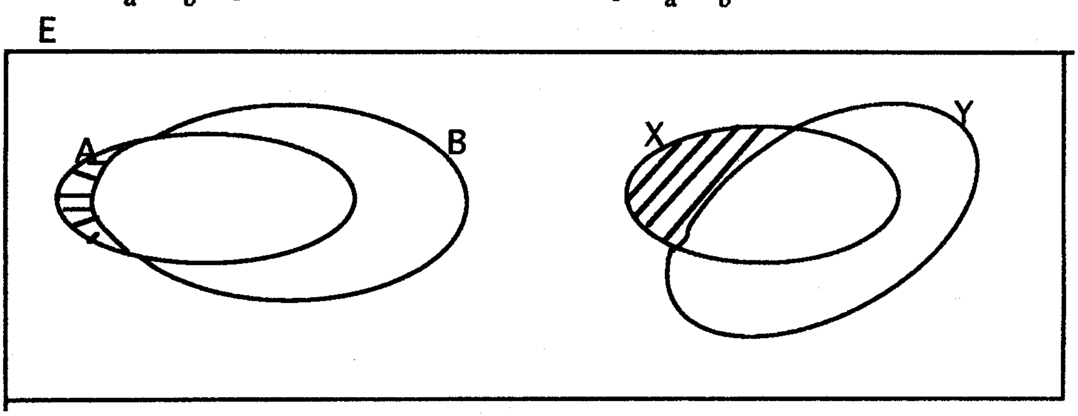
Pour terminer, nous faisons la remarque suivante. Cette résistance liée à la réversibilité peut se rencontrer dans l'apprentissage d'autres notions mathématiques. Il est fréquent, par exemple, d'observer chez des élèves une difficulté énorme lors de l'introduction de la notion de primitives (opération "inverse" de la dérivation) d'une fonction réelle, même en ayant sous les yeux le tableau des dérivées usuelles. On sait d'ailleurs que ce problème de réversibilité explique également la difficulté classique des jeunes élèves lors du passage de l'addition vers la soustraction des nombres.

Annexe : Présentation intuitive de la théorie de l'implication statistique de R.GRAS

L'implication statistique, ou l'analyse implicative, est une méthode classificatoire essentiellement probabiliste d'analyse des données. Elle a la spécificité d'utiliser un indice <u>non symétrique, et sa classification se base sur un concept qui rend compte d'une certaine dynamique orientée au sein de chaque classe constituée.</u>

Considérons par exemple deux attributs **a** et **b** qui sont soumis à **n** individus dont l'ensemble est désigné par E. On ne connaît pas a priori la liaison cognitive qui existe éventuellement entre les deux variables **a** et **b**. Il s'agit de donner un sens statistique à une implication logique et formelle du type " si **a** alors **b**", ((**non a**) ou **b**). Concrètement, on souhaite répondre à la question "*la possession de l'attribut **a** implique-t-elle celle de l'attribut **b**?*" que l'on peut interpréter de diverses manières.

Désignons par A (resp.B) l'ensemble des sujets qui possèdent **a** (resp.**b**). On s'intéresse ici, au lieu de l'intersection, à l'ensemble des sujets qui contredisent l'une des variables, **b** par exemple, dans le cas où les cardinaux n_a et n_b respectivement de A et de B sont tels que $n_a \leq n_b$.



Puisque l'implication **a** ⇒ **b** est contredite sur le sous-ensemble $A\cap\bar{B}$, au lieu de considérer A et B fixés, on les laisse varier aléatoirement et indépendamment dans E en conservant leurs cardinaux respectifs. Comparons alors le cardinal de $A\cap\bar{B}$ à celui qu'aurait l'intersection de 2 parties quelconques X et $\bar{Y}$ de même cardinaux respectifs n_a et $n_{\bar{b}}=n-n_b$ dans une hypothèse d'indépendance ou d'absence de lien entre elles.

Moins il y a d'individus contradicteurs de l'implication, plus l'inclusion qui la représente est statistiquement crédible, donc acceptable. R.GRAS(1979) axiomatise la notion d'implication statistique de la façon suivante :

(a ⇒ b) *est acceptable au niveau de confiance 1-α si et seulement si, dans une hypothèse d'absence de <u>lien a priori (ou d'indépendance)</u>,* $Pr[card(X\cap\bar{Y})\leq card(A\cap\bar{B})]\leq\alpha$.

Ceci signifie que si le le nombre $n_{a\wedge\bar{b}}$ des individus contradicteurs , cardinal de $A\cap\bar{B}$, est "invraisemblablement "petit par rapport à celui qu'on peut attendre de la distribution des cardinaux de $X\cap\bar{Y}$, on acceptera une quasiimplication . La vraisemblance de l'implication **a ⇒ b** sera ainsi mesurée par le complément à 1 de la mesure de la probabilité déduite de la comparaison du cardinal aléatoire de $X \cap \bar{Y}$ à celui de $A \cap \bar{B}$. D'où la définition de la quantité qui mesure l'intensité de l'implication de **a** sur **b** par :

$\varphi(a,\bar{b}) = 1 - Pr[\, Q(a,\bar{b}) \leq q(a,\bar{b})]$ où $q(a,\bar{b})$ est l'indicateur de non-implication , réalisation de la variable aléatoire $Q(a,\bar{b})$ centrée et réduite de la poissonnienne $card(X \cap \bar{Y})$:

soit $Q(a,\bar{b}) = \dfrac{card(X \cap \bar{Y}) - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}}$, et $q(a,\bar{b}) = \dfrac{n_{a \wedge \bar{b}} - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}}$ pour une expérience considérée.

Ainsi, il est clair que les paramètres associés à la définition d'une intensité d'implication entre deux variables données ne dépend que de leurs occurrences et de l'effectif de la population qui les éprouve, mais reste totalement indépendante du nombre des variables restantes. D'ors et déjà, on peut exploiter ce concept d'implication statistique à l'aide de la représentation graphique nommée graphe implicatif, ou réseau de chemins implicatifs, de la relation de préordre partiel $\Re$ induite (à un seuil d'implication fixé (Ψ_0) :

$$\begin{cases} a \,\Re\, b \text{ dès que } \varphi(a,\bar{b}) \geq \Psi_0, \text{et } n_a \leq n_b \\ \text{si } (a \,\Re\, b \text{ et } b \,\Re\, c,\ \varphi(a,\bar{c}) \geq 0{,}5) \text{ alors } a \,\Re\, c \end{cases}$$

est une relation de préordre partiel (donc transitive).

Notons qu'un tel réseau de chemins implicatifs, qui demeure indéfinissable avec un indice de similarité, permet de rendre compte d'une certaine dynamique qui existe entre les attributs descripteurs de ses données. Cependant, on se rend compte que dans le contexte d'une expérimentation effective en vue d'une étude comportementale par exemple, une attitude est rarement isolée ; elle s'accompagne d'autres comportements. Autrement dit, en plus d'une implication entre variables prises deux à deux, son extension en implication entre groupes de variables s'impose. Ainsi, si un groupe de variables est suffisamment désordonné du point de vue implicatif, il n'est pas légitime d'en constituer une classe dont on chercherait l'implication sur une autre classe ou qui serait impliquée par une autre classe. Un tel ordre exigé ou ce désordre rejeté se mesure par exemple par l'entropie de l'expérience, pour chaque couple de variables.

Pour répondre à cette problématique d'extension, A.Larher a élaboré dans sa thèse [26] le concept de cohésion, dit cohésion implicative :

- pour le cas de deux variables a et b , où il y a lieu de parler de l'implication de a sur b, la cohésion implicative du couple orienté (a, b) se mesure par la racine carrée de 1-le carré de l'entropie, si l'implication est supérieure ou égale à 0.5 ; elle est nulle sinon: au désordre maximal, c'est-à-dire quand l'entropie devient maximale (=1), ce qui correspond à une implication égale à 0.5, la cohésion s'annule; elle atteint sa valeur maximale 1 au désordre minimal.

- D'une manière plus générale, pour un groupe A de r variables, la cohésion implicative de A est définie comme la moyenne géométrique des cohésions des couples orientés intervenant : la cohésion implicative d'une classe quelconque $A = \{a_1, a_2, a_3, \ldots, a_r\}$, définie par le couple $A=(((a_1, a_2), a_3) \ldots, a_{r-1}), a_r)$, avec $n_1 \leq n_2 \leq \ldots \leq n_r$ où n_i mesure le nombre d'occurences de a_i, est la quantité

$$\boxed{C(A) = \left[\prod_{i=1,\, j>i,\, j=2}^{i=r-1,\, j=r} c(a_i, a_j)\right]^{2/r(r-1)} .}$$

Conformément à l'intuition, elle s'annulle dès que l'un des couples a sa cohésion nulle.

Enfin, l'implication entre deux classes est définie comme une quantité qui rend compte de l'intensité d'implication maximale des éléments d'une classe sur les éléments de l'autre et de la valeur des cohésions des classes intervenant par leur moyenne géométrique :

Soit $A=\{a_1, a_2, \ldots, a_r\}$ et $B=\{b_1, b_2, \ldots, b_s\}$ deux classes (disjointes) d'attributs, ces derniers étant soumis à une même population de sujets.

Conformément aux lois de probabilité des sup. des variables aléatoires, l'**indice d'implication** $\Psi(A,B)$ **de la** **<u>classe A sur la classe B est définie par la quantité :</u>**

$$\boxed{\Psi(A, B) = [\sup_{i,j} \varphi(a_i, \bar{b}_j)]^{rs} \times [C(A) \times C(B)]^{1/2}, \text{ avec } i‘\{1, \ldots, r\} \text{ et } j‘\{1, \ldots, s\}.}$$

Pour de plus amples informations concernant les propriétés et la justification de cette définition , nous renvoyons à la thèse de A.LARHER sus-citée, pages 40, 41 et 42.

Remarquons que l'implication entre classes devient nulle dès que l'une des classes l'est déjà. Elle décroît avec les effectifs des classes, mais demeure une fonction croissante de la cohésion de chaque classe.

Dès lors, il est devenu possible de représenter la structure existant entre les variables descripteurs à l'aide d'un dendogramme construit par agrégations successives en commençant par le couple de variables ou de classes de variables déjà construites de cohésion maximale à un niveau donné : c'est l'arbre de classification hiérarchique implicative et cohésitive (A.C.H.I.C.) ou tout simplement arbre implicatif.

<u>Pratique :</u> Il y a trois types possibles de variables traitables : - variables booléennes, ou présence-absence (cf. exemple introductif du présent annexe, et les 10 variables considérées dans VIII -B.1 & 2);

- variables à nuances, ou modales ou ordinales ; - variables quantitatives ou numériques.

Si l'on s'intéresse à l'étude des liaisons implicatives entre variables, il faut construire et interpréter un graphe implicatif des descripteurs, à un seuil d'implication fixé. Si l'on vise l'étude des liaisons implicatives entre groupes des variables, il faut interpréter les diverses sous-implications sur l'A.C.H.I.C.

Références et bibliographie

[1] **M.ARTIGUE**, 1991, *Epistémologie et didactique, in* RDM Vol.10/2.3. p.249.

[2] **I.M.BERONDO-AGRE & Per AGRELL**, 1992, *Vers une syntaxe des diagrammes de Venn : lutte contre un mythe, in* Journal de la Société de Statistique de Paris n° 1/2-1er et 2è trimestre 1992 (pp 134-141).

[3] **C.BLOCH**, 1974, Elément de réponses à une question concernant les programmes (1ère & terminale) - *"Faut-il enseigner une axiomatique des Probabilités et laquelle?", in* L'Enseignement des Probabilités et des statistiques. Commission Internationale pour l'Etude et l'Amélioration de l'enseignement des Mathématiques. CR de la 26ème rencontre. Bordeaux aout 1974.pp.139-161. IREM de Bordeaux.

[4] **A.BODIN**, 1992, *Réflexions sur les représentations, les conceptions et les compétences à partir d'une évaluation à grande échelle des programmes de mathématiques de l'Enseignement secondaire (EVAPM).* I.R.E.M. de Besançon.

[5] **K. BOGNARY & T.,NEMETZ** 1977, *On the teaching of probability at secondary level.* in Educational Studies in Mathematics 8(1977) pp.399-404.

[6] **J.BORDIER**, 1991, *Un modèle didactique, utilisant la simulation sur ordinateur pour l'Enseignement de la probabilité.* Thèse présentée à l'université Paris VII.

[19] **G.BROUSSEAU & J.BRIAND**, 1974, *Généralités sur l'Enseignement des Probabilités au niveau élémentaire, in* L'Enseignement des Probabilités et des statistiques. Commission Internationale pour l'Etude et l'Amélioration de l'enseignement des Mathématiques. CR de la 26ème rencontre. Bordeaux aout 1974. pp.66-123. IREM de Bordeaux.

[7] **G.BROUSSEAU**, 1986, *Théorisation des phénomènes d'enseignement des mathématiques.* Thèse de Doctorat ès Sciences. Université de Bordeaux I.

[8] **G.BROUSSEAU**, 1989, *Les obstacles épistémologiques et la didactique des mathématiques, in* Construction de Savoir. Obstacles & Savoirs. CIRADE. pp.41-63.

[9] **G.BROUSSEAU**, 1983, *Les obstacles épistémologiques et les problèmes en mathématiques,* RDM vol.4.2, pp.165-198.

[10] **R.DOUADY** 1984, *Jeux de cadres et dialectique objet-outil dans l'enseignement des mathématiques.* Thèse de Doctorat ès Sciences. Université Paris VII.

[11]**R. DUVAL** 1988, *Ecarts sémantiques et cohérence mathématique : introduction aux problèmes de congruence, in* Annale des didactique et et sciences cognitives, *in* annales de didactiques de mathématiques, 1, IREM de Strasbourg.

[12] **R. DUVAL & MA.EGRET**, 1990, Pour une décomposition cognitive des tâches de production d'une démonstration, *in* Publication de l'IRMAR, Fascicule 5 (1989/1990).

[13] **A.ENGEL**, 1974, *Les abaques probabilistes, in* L'Enseignement des Probabilités et des statistiques. Commission Internationale pour l'Etude et l'Amélioration de l'enseignement des Mathématiques. CR de la 26ème rencontre. Bordeaux aout 1974. pp.2-25. IREM de Bordeaux.

[14] **R.FALK - R.FALK - I.LEVIN**, 1980, *A potentiel for learning probability in young children.* Educational Studies in Mathematics. Vol.11, n°2, pp.181-204.

[15] **FELLER William**, 1950, *Probability theory.* Vol.1. John Willey. New York.

[16] **E.,FISCHBEIN, MARIA SANITA NELLO & MARIA SCIOLIS MARINO**, 1991, *Factors affecting probabilistic judgements in children and adolescents.* Educational Studies in mathematics. Vol 22, pp.523-549.

[17] **M.FLATO**, 1990, *Question de science : Le pouvoir des Mathématiques.* Hachette.

[18] **B.V. GNEDENKO & A.Ia.KINTCHINE**, 1960, *ntroduction à la théorie des probabilités.* Monographie Dunod.

[19] **R.GRAS**, 1979, *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et certains objectifs didactiques en mathématiques*. Thèse de Doctorat ès Sciences. Université de Rennes I.

[20] **R.GRAS**, 1991, *L'analyse des données : une méthodologie de traitement de questions didactiques*. Actes de la VIème Ecole et Université d'été de didactique des Mathématiques 1991 et RDM 12/1, 1992.

[21] **R.GRAS & A.LARHER**, 1992, *L'implication statistique, une nouvelle méthode d'analyse des données, in* Mathématique, Informatique et Sciences Humaines n°120 (à paraître).

[22] **P.L.HENNEQUIN**, 1991, *Quelques éléments pour un débat sur la place de la statistique et du calcul des probabilités dans une formation de base de niveau universitaire*. Gazette des Mathématiciens . Avril 1991/Supplément n°48.

[23] **P.L.HENNEQUIN** , 1990, *Indépendance et indépendance conditionnelle*. Bull.APMEP n°376.

[24] **G. I.KEMENY.- J.L.-SNELL G.L. THOMPSON** , *Algébre Moderne et Actvités humaines*, vol.7, Dunod, 1969, pp. 130-131.

[25] **A.N.KOLMOGOROV** , 1933, *Fondation s of the theory of tprobability* TChesla publishing company New York(1956)

[26] **A.LARHER**, 1991, *Implication statistique et applications à l'analyse de démarche de preuve mathématique*. Thèse de Doctorat d'Université. Université de Rennes I. U.F.R. de Mathématiques.

[27] **S.MAURY**, 1986, *Contribution à l'étude didactique de quelques notions de probabilité et de combinatoire à travers la résolution de problèmes*. Thèse de Doctorat d'Etat. Université des sciences et Techniques du Languedoc.Montpellier.

[28] **A.MESQUITA**, 1989, *L'influence des aspects figuratifs dans l'argumentation des élèves en géométrie : élément pour une typologie*. Thèse de Doctorat de l'Université Louis Pasteur. Publication de l'Institut de Recherche Mathématique Avancée de Strasbourg.

[29] **B.PARZYS** , 1990, *Un outil sous-estimé : l'arbre probabiliste*. Bulletin APMEP n°372.

[30] **J.PIAGET & B.INHELDER** , 1951, *La genèse de l'idée de hasard chez l'enfant*. PUF Paris.

[30'] **F.PLUVINAGE**, 1977, *difficulté des exercices scolaires en mathématiques*. Thèse de doctorat ès sciences. ULP Strasbourg.

[31] **H.POINCARE**, 1912, *Calculs des probabilités*. Cours de la faculté des sciences. Gauthier-Villars.

[32] **J.M.SHAUGHENESSY** , 1977, *Misconceptions of probability : an experiment with a small-group, activity-based, model building approach to introductory probability at the college level*. Educational Studies in Mathematics. Vol 8, n°3, PP.295-316.

[33] **H.STEINBRING** , 1986, *L'indépendance stochastique*, R.D.M., Vol.7, n°3, 5-50.

[34] **TERRENCE L. FINE**, 1973, *Theories of probability . An examination of foundations*. Academic press.

[35] **A.TOTOHASINA**, 1992, *Conceptions causaliste ou chronologiste de la notion de probabilité conditionnelle, in* Cahier de didactique fasc.9. Departement de Mathématique de l'institut français de Tessalonique (original en français accompagné de sa traduction en grec).

[36] **A.TOTOHASINA**, 1992, *Méthode implicative en analyse des données et application à l'analyse des conceptions d'étudiants sur la notion de probabilité conditionnelle*. Thèse de doctorat. Université de Rennes I.

[37] **G.VERGNAUD**,1981, *Quelques orientations théoriques et méthodologiques des recherches françaises en didactique des mathématiques*. R.D.M. Vol.2/2, pp.215-232.

[38] **G.VERGNAUD**,1988, *Question de représentation et de formulation dans la résolution de problèmes de*

mathématiques, in Annales de didactique et de sciences cognitives 1(1988) pp.33-35. IREM de Strasbourg.

[39] **G.VERGNAUD**,1990, *La théorie des champs conceptuels.* R.D.M. Vol.10/2.3. Ed. La pensée sauvage. Grenoble.

[40] **G.VERGNAUD**,1991, *Morphismes fondamentaux dans les processus de conceptualisation, in* Les sciences cognitives en débat; Editions du CNRS, Paris.pp.11-28.

[41]**G.VERGNAUD**,1989, *Difficultés conceptuelles, erreurs didactiques et vrais obstacles épistémologiques dans l'apprentissage des mathématiques, in* Construction des Savoirs. Obstacles & Conflits. CIRADE. pp.33-40.

[42] **T.H.WONNACOTT & R.J.WONNACOTT**, 1984, *Statistique, Economie-Gestion-Sciences-Médecine.* Economica. 3è édition. pp.75-87.

[43] **M.ZAKI**, 1990, *Traitements de problèmes de probabilités en situation de simulation.* Thèse de Doctorat d'université. Publication de l'I.R.M.A. de l'Université Louis Pasteur. Strasbourg.