# L'INTRODUCTION DU CONCEPT DE PROBABILITE CONDITIONNELLE : AVANTAGES ET INCONVENIENTS DE L'ARBORESCENCE

André TOTOHASINA
IRMAR de Rennes

Le présent article se centre essentiellement sur la notion de probabilité conditionnelle, discipline d'enseignement. Rappelons le concept. Soit $(\Omega, \beta, p)$ un espace probabilisé, ($\Omega$ est l'ensemble fondamental, $\beta$ une tribu sur $\Omega$ et p probabilité définie sur $\beta$) et A un élément de $\beta$, de probabilité non nulle. La *probabilité conditionnelle sachant A* est l'application de $\beta$ dans [0, 1], qui à tout événement B de $\beta$ associe le réel $p_A(B) = \dfrac{p(B \cap A)}{p(A)}$. On démontre que $(\Omega, \beta, p_A)$ est également une probabilité. Une probabilité conditionnelle est généralement différente de la probabilité initiale. Il est clair que c'est le concept qui donne son sens au symbole, et non l'inverse comme le font constater parfois certaines productions d'élèves en premier apprentissage de la probabilité conditionnelle selon la notation utilisée. Une étude didactique suffisamment profonde PARZYSZ [22] conduirait à adopter la notation indicielle lors de sa première introduction. En effet, l'événement conditionnant joue un rôle qui serait plus rapprochable à celui joué par un paramètre que par une variable. La notation $p(B/A)$ pourrait donner aux débutants l'impression d'une fonction p appliquée à deux variables B et A. Quant au mode de lecture, adopter la locution *probabilité sachant A de (l'événementt) B*, de la même manière que *logarithme de base a de x* pour $log_a(x)$, par exemple, semblerait moins dépaysant aux débutants. 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 [28], les conceptions causaliste et chronologiste, sur lesquelles peuvent se fonder de véritables obstacles épistémologiques [7] :

— *conception chronologiste* d'une probabilité conditionnelle : c'est le fait d'impliquer systématiquement la "chronologie " ou le "temps" dans la façon d'appréhender ce concept ;

— *conception causaliste* 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 ou soit tout au moins précaire 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". R.FALK (1989) mentionne également ces deux difficultés et propose une situation concrète qui aiderait entre autres les élèves à surmonter la résistance liée à cette conception chronologiste. L'idée de la situation proposée est la suivante. On considère quatre boules identiques, indiscernables au toucher, et dont deux sont noires et les autres blanches, contenues dans une urne opaque. Après avoir agité l'urne, on en tire une boule que l'on maintient cachée dans sa poche. De la même urne, on tire ensuite une deuxième boule que l'on montre cette fois à la classe entière : elle est blanche. On demande alors d'évaluer la probabilité que la boule gardée dans la poche soit blanche aussi. Cette situation demeure naturellement très modulable selon le niveau des élèves. 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 [29] : c'est le fait de transporter systématiquement la quantification traditionnelle d'une probabilité correspondant à une hypothèse d'équiprobabilité d'événements élémentaires : "cas favorables sur cas possibles" à la probabilité conditionnelle ; cela se concrétise par la tendance systématique à se représenter la *probabilité conditionnelle de B sachant A* : $p_A(B)$, par le rapport $p_A(B) = \dfrac{\text{Card}(B \cap A)}{\text{Card}A}$

ou , à tort, par $p_A(B) = \dfrac{\text{Card}B}{\text{Card}A}$ . Elle *s'oppose donc à la conception fondamentaliste* qui consiste à utiliser la définition formelle de ce concept. On pourrait évoquer d'autres obstacles (*cf.* par exemple [24a] avec ses analyses approfondies et suivies des suggestions intéressantes sur les plans pédagogique et recherche). Entre autres, il y a la *conception subjectiviste* renforcée par l'idée qu'une information nouvelle apportée au sujet modifie la probabilité de l'événement. A mon sens, les deux premières conceptions relèvent du cognitif, et la troisième de l'opératoire ou de la transposition didactique. Dans le présent article, je propose une méthode pédagogique qui tente justement de contrôler ces trois représentations de la notion de probabilité conditionnelle.

## I. 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* [1] *stochastique* et celle d'*incompatibilité* d'événements, le comportement *cardinaliste*, etc. Par contre, la formule des 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 problème qui mettrait en œuvre la formule de Bayes (d'une manière implicite, car elle n'est pas au programme du secondaire) sans que l'élève ait besoin de mobiliser les prérequis supposés nécessaires (entre autres, nous pensons à l'acquisition du vocabulaire élémentaire des ensembles : intersection, réunion etc). Alors comment faire ? Ma première réponse s'oriente vers l'utilisation d'un outil graphique (en l'occurrence un arbre des probabilités), permettant de sortir du cadre 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. Pour cela, je renvoie à l'article de M. HENRY [16].

(1) 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.

## II. 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 XVIII$^e$ siècle nous confirme 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 1/4 ; 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. Le raisonnement ainsi présenté décrit un mode de calcul sur un exemple concret qui a été ensuite formalisé et généralisé. Il est vrai que $p_C$ était à l'époque admise comme probabilité, mais essentiellement de façon opératoire. Par ailleurs, la locution comme "*1/4 de 1/10*", qui rappelle la notion de partage, pour exprimer 1/4 × 1/10 paraît également très intéressante; 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 ([13], p. 21). Mais, puisqu'on n'a pas dégagé formellement la généralité du raisonnement, tout en étant conscient de sa portée puisqu'on le prend pour exemple explicite : c'est le contraire d'un théorème en acte. En revanche, on pourrait atteindre ce dernier. D'ailleurs, certains auteurs, comme B.V. Gnedenko & A.L. Khintchine [12], 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 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, en vue de préparer aux séquences didactiques, je suis amené à conjecturer, à des fins didactiques, 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 propose aux élèves d'interpréter provisoirement P(A) par "pourcentage de chance de réalisation de A". On terminera la 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), *obtenue à partir d'un exemple, on tire* $P(B \mid A) = \dfrac{P(A \text{ et } B)}{P(A)}$ . 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 par rapport à la probabilité conditionnelle ?*

— **b**) l'analyse arborescente (c'est-à-dire l'utilisation d'un arbre des probabilités (*cf.* B. PARZYSZ [23])) permettrait de sauter sans trop de risques l'étape des opérations ensemblistes, comme on l'a déjà signalé plus haut, en l'occurrence la distributivité de l'*intersection* par rapport à l'*union*. 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 le problème suivant sur lequel un stagiaire est venu discuter :

> *Soit un jeu à deux issues : gagner et perdre. 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 + 1$^{\text{ème}}$ coup, et B au 2k $^{\text{ème}}$ coup... Donc, par exemple, l'une des probabilités demandées est :

$$p(\{A \text{ gagne}\}) = \sum_{k=0}^{+\infty} p(\{A \text{ gagne au } 2k+1^{\text{ème}} \text{ coup}\})$$

$$= p_1 \sum_{k=0}^{+\infty} (1-p_1)^k (1-p_2)^k = \frac{p_1}{1-(1-p_1)(1-p_2)} ;$$

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

Alors j'ai produit le graphe ci-dessous :

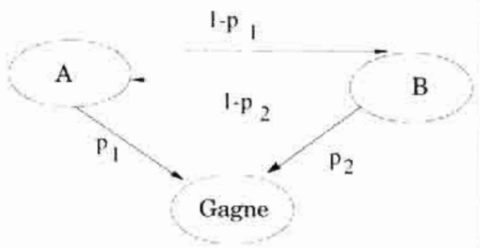


**Fig. 1**

Ensuite, j'ai présenté l'arbre des probabilités de la figure 2 :

"...Durant les 9 premiers coups, on a : (A gagne au 1er coup), ou (A perd au 1er coup) et (B perd au 2e coup) et (A gagne au 3e coup), ou..., ou (A gagne aux 1er, 3e, 5e et 7e coups) et (B perd aux 2e, 4e, 6e et 8e coups)..." Enfin, le stagiaire se 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 [21]. 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 le raisonnement sur un graphe. En plus de sa *richesse* en information, cela permettrait aussi de gagner énormément de temps par rapport à d'autres méthodes algébriques. Enfin, une fois cette représentation, si l'on peut dire, "maîtrisée" dans le cas de deux "conditions" et deux "produits", il serait envisageable de généraliser à 3, 4, etc.

## III. Méthodologie sur l'action didactique effective

Je propose donc de :

* présenter un problème qui ne contient que des pourcentages et qui ne présuppose pas l'utilisation d'une probabilité uniforme, afin d'éviter surtout une éventuelle complication d'ordre combinatoire; une situation qui de plus prête peu à une causalité et à une conception chronologiste, ni d'une manière évidente, à un comportement cardinaliste,

* faire émerger et définir l'algorithme de calcul à partir d'un exemple,

* donner enfin une définition formelle d'une probabilité conditionnelle,

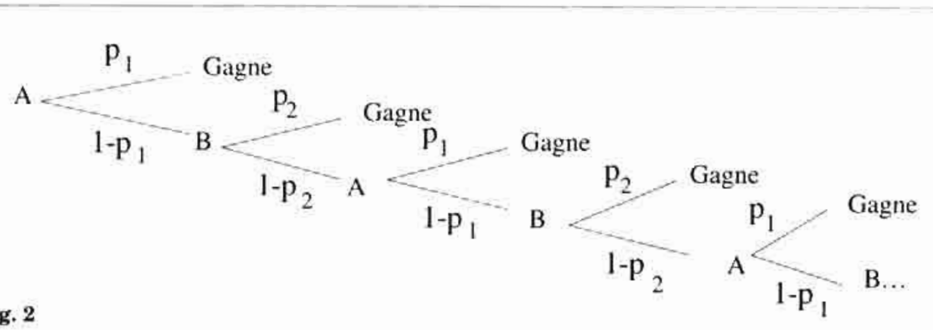


**Fig. 2**

* 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 représentation engagée une méthode pour traiter une classe de situations-problèmes plus ou moins proches de 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 chance assez forte pour faciliter l'intériorisation, c'est-à-dire l'appropriation par une majorité d'élèves, de la notion en question.

## IV. Un exemple de problème

### IV.- 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 boîte rouge".*
*B : "obtenir un carton contenant une boîte bleue".*

### IV.- B. Gestion du déroulement de la séquence et analyse *a priori*

### IV.- B.1. Plan

* Présenter le problème, sans avoir mis le titre du chapitre, objet de l'étude, "Probabilité conditionnelle". Cela devrait éviter une trop grande directivité. 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 au cas où elles seraient erronées.

* Poser des questions 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.

* Donner la définition axiomatique d'une probabilité conditionnelle seulement après une petite synthèse du problème traité.

* Prévoir un test après le cours.

### IV.-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 aussi à d'autres propositions, par exemple $\Omega$ ={bleue, rouge, marque M} qui n'est pas pertinente évidemment, car les événements "contenir une boite 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 porteraient 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 $(\Omega, \beta, P)$ avec, comme univers $\Omega = \Omega_1 \times \Omega_2$, où $\Omega_1 = \{R, B\}$ et $\Omega_2 = \{M, \overline{M}\}$, comme tribu $\beta = \mathcal{P}(\Omega_1) \times \mathcal{P}(\Omega_2)$, comme probabilité $P = p_1 \otimes p_2$, où $p_i$ 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 introduite sur les tirages pourrait facilement créer une résistance chez des élèves débutants relativement à la *notion nouvelle* de probabilité conditionnelle.

**— 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,* et non *"on sait que 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 qu'ils vont proposer. En tout cas, le problème est ainsi *dévolu* à la classe qui doit vivre une sorte d'amorce d'un *échange intellectuel*. ( 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. Il faut aboutir à ce que les élèves considèrent qu'il s'agit d'une autre probabilité. 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 d'atténuer sa chance de répétition. Cela ne suffit pas à lever l'obstacle, car 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 habituel. 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 :* arbre des probabilités.** On peut faire allusion à un "arbre de dénombrement ", technique traditionnellement déjà pratiquée dans le chapitre sur le dénombrement, éventuellement 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 ce problème à l'aide d'un arbre :

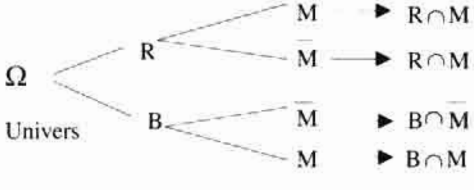


**Fig. 3**

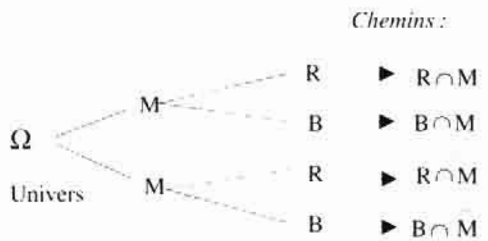


**Fig. 4**

*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 3 ci-dessus (voir PARZYSZ [23] ) :

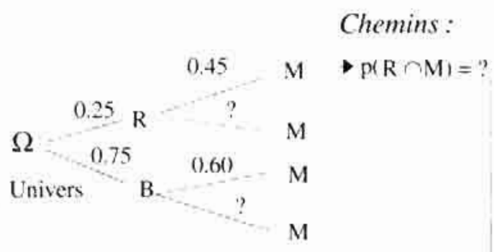


**Fig. 5**

Dans les conventions, il y a le produit $p(M \cap R) = 0.45 \times 0.25$ qui signifie ici 45% de 25%. On est confronté ainsi au problème des arbres, outils de représentation et non de justification des fondements. Vu les conventions précédentes, les élèves ne peuvent pas faire autrement que proposer : $P_R(M) = 0.45$, ou encore $P_R(M) = \dfrac{45}{100} = \dfrac{9}{20}$ . Sinon, quel sens donneraient-ils à R — M ? Je suggère, à cette occasion, d'amorcer le concept de conditionnement : la probabilité de l'événement M, liée à la condition "le carton contient une boite 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 : $p_B(M)$, $p_B(\overline{M})$ 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 complète 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 :

$$p(R) = \frac{25}{100} = \frac{1}{4} \quad \text{et} \quad p_R(M) = \frac{45}{100} = \frac{9}{20} \; .$$

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) Deuxième question :*

*Quelle est la proportion des cartons (R et M) dans le magasin ?*

La réponse attendue est :

$$\frac{9}{20} \text{ de } \frac{1}{4} = \frac{9}{20} \times \frac{1}{4} = \frac{9}{80} \quad .$$

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* [29] 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) Troisième question :*

*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 :

$$p(M \cap R) = \frac{9}{80} = p_R(M) \cdot p(R) \quad .$$

La notion de probabilité conditionnelle apparaît ainsi comme un outil pour le calcul de la probabilité conjointe $p(M \cap R)$. 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 :

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

Précisons qu'il ne s'agit pas ici d'une institutionnalisation, mais d'une simple traduction schématique d'une situation mathématique.

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 :

$$p_R(M) = \frac{p(M \cap R)}{p(R)} \quad .$$

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 limiter

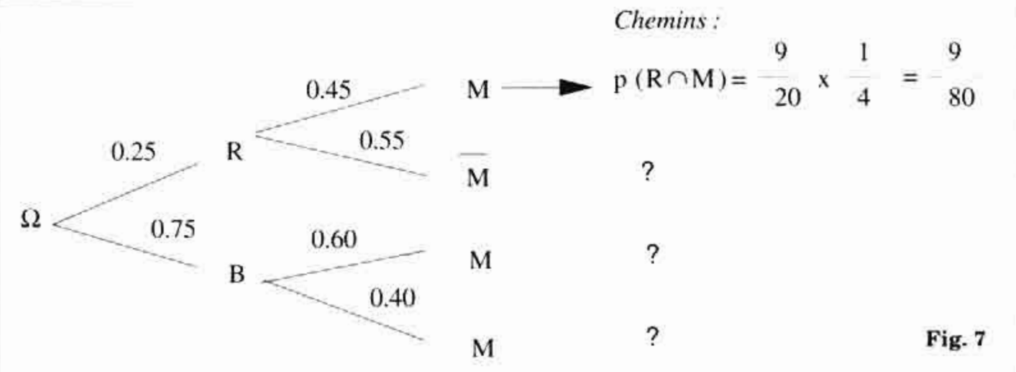


**Fig. 7**

un éventuel *comportement cardinaliste*[2] qui risque de devenir réducteur, chez les élèves. Je renvoie à M.HENRY [17]) qui parle de l'introduction fréquentiste de probabilité.

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 7 ci-dessus (qui reste encore à compléter évidemment).

**Pour la question Q3 :** détermination de p(M)

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

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

(2) Il s'agit du comportement qui consiste à croire qu'une probabilité conditionelle $p_R(M)$ 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. [29]. chap. III).

Ensuite, dans le souci de renforcer la compréhension du mécanisme de calcul ainsi pratiqué, il est loisible de demander à la classe de 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 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) :

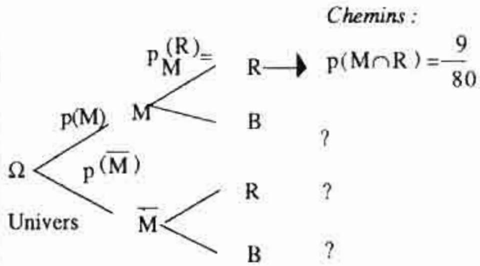


**Fig. 8**

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, d'ailleurs connotée d'une légère sémantique chronologique. Cette fois, la classe sera nécessairement amenée à résoudre en x des équations du type $p(M \cap R) = x.p(M)$ ; soit

$$x = p_M(R) = \frac{p(M \cap R)}{p(M)}$$

Comme exercice d'entraînement, on pourra ensuite demander de calculer de deux façon les probabilités $p_R(B)$, $p_{\overline{M}}(R)$, 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, avec p(M) et p(R) non nulles, on a les relations du type :

$$p(M \cap R) = p(R \cap M)$$

$$p_R(M) \times p(R) = p_M(R) \times p(M)$$

Ce qui permet d'avoir :

$$p_R(M) = \frac{p(M \cap R)}{p(R)} = \frac{p_M(R) \times p(M)}{p(M \cap R) + p(R \cap \overline{M})}$$

si l'on connaît $p_M(R)$ et p(M) ; ou

$$p_M(R) = \frac{p(M \cap R)}{p(M)} = \frac{p_R(M) \times p(R)}{p(M \cap R) + p(M \cap \overline{R})}$$

si l'on connaît $p_R(M)$ 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 $p : \mathcal{P}(\Omega) \to [0, 1] : A \to p(A)$ 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 entraîne que $p_B$ est bien une probabilité, et différente de p. On peut noter que si A et C sont tels que $p(A) \neq p(C)$ et $A \cap B = C \cap B$ alors $p_B(A) = p_B(C)$.

3) Faire vérifier que $p_B$ est effectivement une probabilité. 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** $\Omega$ (c'est aussi une probabilité sur B) :

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 [1], on admettra que $B \cap (A \cup C) = (B \cap A) \cup (B \cap C)$, avec $(B \cap A)$ et $(B \cap C)$ incompatibles. $P_B(A \cup C) = P(B \cap (A \cup C))/P(B) = P((B \cap A) \cup (B \cap C))/P(B) = [P((B \cap A) + P(B \cap C)]/P(B)$ ; soit $P_B(A \cup C) = P_B(A) + P_B(C)$. En fait, une telle vérification (de niveau plutôt universitaire) n'est pas nécessaire au niveau secondaire, me semble-t-il, $\Omega$ étant constitué d'événements élémentaires en nombre fini $\omega_1, \dots, \omega_n$ : il suffit que

$$\sum_{i=1}^{i=n} P_B(\omega_i) = 1 \ .$$

*Note :* On admettra alors que $p_B$ satisfait les axiomes de probabilité, et donc qu'elle en possède toutes les propriétés. *En particulier :* On vérifie que : $p_B(A) = 1 - p_B(A)$ ; $p_B(\varnothing) = 0$ ; $p_B(A \cup C) = p_B(A) + p_B(C) - p_B(A \cap C)$. Puis, 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 $p(A \mid B)$, en indiquant l'éventuelle conséquence infondée de croire que (A/B), ou (A | B) sont des événements (3), pourtant le pli est pris en terminale.

Enfin, terminer en traitant un exemple simple, mais qui n'est pas issu d'un même modèle que celui 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.

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

(3) Dans notre étude 29] 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.

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 1re boule tirée est rouge " ; R2 : "la 2e 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,

$$p_{R1}(R2) = \frac{p(R1 \cap R2)}{p(R1)} = \frac{p(R)}{p(R1)} \quad .$$

Or $p(R) = 3/10$ et $p(R1) = 3/5$. D'où

$$p_{R1}(R2) = \frac{3/10}{3/5} = \frac{1}{2} \quad .$$

On termine ensuite par le renforcement suivant pour le calcul de $p(R)$ :

$$\begin{aligned} p(R) = p(R1 \cap R2) &= p_{R2}(R1) \times p(R2) \\ &= p_{R1}(R2) \times p(R1) \text{ (de préférence ici...)} \end{aligned}$$

Je renvoie à l'exemple des jetons dans M. HENRY (1993).

## V. Bilan qualitatif d'une expérimentation

*L'expérimentation faite auprès des élèves d'une classe de terminale D* (4) *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 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 ; mais, cela mérite une reconfirmation car d'autres résultats montrent le contraire ; on se demande également de ce qu'il est de la notion de proportion ?

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 semble offrir 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 ; il n'empêche qu'elle est expressément exclue du programme, car cela demande un coût épistémologique que de contredire trop tôt l'approche fréquentiste, à moins d'une précaution particulière ;

4) aucun indice de conception causaliste de cette notion ne s'est manifesté ;

(4) Cette classe est tenue sous la responsabilité de M. A.Simon un collègue de recherche, membre de l'équipe didactique de l'I.R.MA.R. Je le 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.

5) par rapport à d'autres méthodes d'introduction de la notion de probabilité conditionnelle (motivation par fréquences conditionnelles (M. HENRY [17]...), 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, la remarque la plus importante 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. D'où la nécessité de présenter plusieurs approches et de passer de l'une à l'autre. Par exemple, il nous semble maintenant utile d'alterner avec un autre mode de représentation de la situation et celui du niveau des conceptions. Pour être clair, il serait possible d'indiquer, concurremment à l'arbre, une représentation à l'aide d'un "tableau à double entrée",

| | M | $\overline{M}$ | |
|---|---|---|---|
| R | | | $\Omega$ |
| $\overline{R}$ | | | |

ou tout simplement à l'aide d'une "patate" comme dans la pratique traditionnelle. On verra plus tard les difficultés introduites 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. Il faudrait lancer une expérimentation à plus grande échelle.

## VI. Evaluation immédiate : le lendemain de la séance du cours

### VI.-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'un problème isomorphe à celui traité 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 "familier" 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. On s'attend également à l'écriture formelle de la probabilité cherchée, avant calcul numérique

## VI.- 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 :

$$p_B(U_1) = \frac{\mathrm{card}(U_1 \cap B)}{\mathrm{card}(B)} \quad ,$$

formule qui n'est valable que sous hypothèse d'équiprobabilité des urnes $U_1$ et $U_2$.

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

$$p_B(U_1) = \frac{p(U_1 \cap B)}{p(B)} = \frac{p_{U_1}(B) \cdot p(U_1)}{p(B)} \quad ,$$

$$\text{avec } p(B) = p(B \cap U_1) + p(B \cap U_2) \quad ;$$

$$\text{or } p(B) = p_{U_1}(B) \cdot p(U_1) + p_{U_2}(B) \cdot p(U_2)$$

$$= \frac{1}{2} \times \frac{2}{5} + \frac{1}{2} \times \frac{4}{5} = \frac{3}{5} \quad ;$$

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

$$p_B(U_1) = \frac{1}{3} \approx 0.33 \quad .$$

Enfin, chacune de ces deux procédures de résolution pourrait également s'accompagner d'une analyse arborescente du genre ci-dessous.

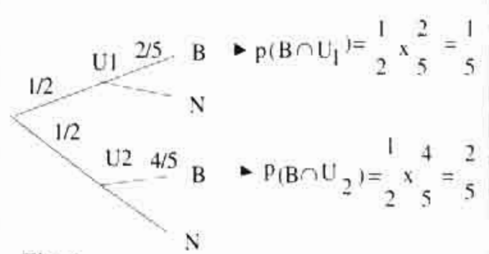


**Fig. 9**

## VII Analyse des démarches effectives de réponses des élèves (5)

**Sommaire :** ces deux exercices sont proposés également aux élèves de la terminale A1 du Lycée Ile-de-France (professeur Madame A. Larher) qui venaient de recevoir un enseignement sur la notion de probabilité. Mais la situation didactique menée dans cette classe ne pouvait pas mobiliser l'arborescence en tant que connaisance institutionnalisée contrairement à la terminale D précédente où se passait l'expérimentation véritable. Pour indiquer brièvement une comparaison entre les performances globales des deux classes, voir le tableau de réussites-échecs relatifs à ces deux situations et aux deux classes (figure 10).

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

**Fig. 10**

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

### Analyse statistique de quelques croisements de réussites-échecs

### Classe TD : Tableau de Réussites-Echecs aux deux problèmes du premier test

| Problème d'urnes | Fabrication des boulons: Réussite | Fabrication des boulons: Echec | |
|---|---|---|---|
| Réussite | 12 | **0** | 37.5% 12 |
| Echec | **8** | 12 | 62.5% 20 |
| | 20 62.5% | 12 37.5% | 32 |

**Fig. 11**

Le tableau ci-dessus suggère qu' il y aurait donc *dépendance entre les résultats à ces deux problèmes.* Vraisemblablement, le

(5) Le mot "réussite" signifiera que le problème est résolu correctement par l'élève. Le mot "Echec" s'y oppose.

"problème d'urnes" apparaît plus difficile que l'autre problème, le taux de réussite étant plus faible. 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 |

**Fig. 12**

Ces deux tableaux (fig. 12 et 13) nous indiquent une légère diminution de la pratique de l'analyse arborescente pour le problème d'urnes où ne figurent plus des pourcentages (6). 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

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

**Fig. 13**

figurant dans l'énoncé du "problème d'urnes" à des pourcentages (*cf.* les exemples de productions écrites relevées dans [29] chap. V).

### Quelques éléments d'explication sur les démarches non pertinentes

Je me contenterai ici de l'évaluation 1 (Pour la consultation de quelques copies d'élèves, voir [35], 239-260). Grosso modo les erreurs les plus marquantes portent sur deux points :

**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. Par exemple, à propos du "problème de boulons", examinons les deux réponses proposées par deux élèves de terminale A1 ci-dessous :

(6) Pour le 1er tableau (fig.12) : $\chi^2_{\text{Mac-Némar corrigé}} = 6.7$, et pour le 2e (fig. 13) : $\chi^2_{\text{Mac-Némar corrigé}} = 9.6$.

***Réponse 1** :Une telle réponse témoigne d'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 la ligne "boulons défectueux " avec la colonne "M1" laisse penser qu'il y a la confusion entre probabilité conditionnelle et probabilité conjointe. Pour cet élève, ce mode de représentation figurale en tableau ne semble pas bien refléter le conditionnement d'événément.*

| | $M_1$ | $M_2$ |
|---|---|---|
| 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 defectueux " $p(\bar{B}) = \frac{6}{100}$

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

***Réponse 2 :***

$M_1$ fabrique 0,4 de la totalité des boulons

40+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 |

Défectueux = D

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

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

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-congruent*e (7) 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". 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 R. DUVAL [8]) de notations symboliques (ou formalisme), le registre langagier mathématique, le registre langagier naturel, le registre figural. A. MESQUITA a montré dans sa thèse [20] 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 [20] *(ibid)*.

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 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 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

(7) Ce terme de congruence (ou de non-congruence) sémantique est utilisé ici au sens de R. Duval (1988) et de A. Mesquita (Thèse, 1988, chapitre 1er) à propos de l'apprentissage de la démonstration mathématique chez des jeunes élèves.

notion de probabilité conditionnelle, mais risque d'ériger un obstacle dans les problèmes bayesiens et renforce l'aspect temporel ou successif que les programmes voulaient exclure. En conséquence, une prudente réserve didactique s'impose. Cette notion de congruence sémantique, semble expliquer en partie une conclusion de S.MAURY[19] 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.* 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 [31] 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 qui est priviligiée (chaque case représentant une intersection). Un exemple d'homomorphisme est proposé dans Totohasina ([29], 237-238).

**Obstacles conceptuels :** ils se situent essentiellement à trois niveaux :

— l'évaluation de la probabilité de l'événement conditionnant et la reconnaissance de l'expérience aléatoire dont il est issu ;

— 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 au 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) que 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*. "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"(R. GRAS ([13], 15-16)). Ainsi, on peut penser que 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.

## REFERENCES ET BIBLIOGRAPHIE

[1] 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 2e trimestres 1992, 134-141.

[2] C. BLOCH, 1974, "Elément de réponses à une question concernant les programmes (1re & 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 26e rencontre. Bordeaux août 1974, IREM de Bordeaux., 139-161.

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

[4] J. BORDIER, 1991, *Un modèle didactique, utilisant la simulation sur ordinateur pour l'Enseignement de la probabilité.* Thèse de Doctotorat, Université Paris VII.

[5] 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 26e rencontre, Bordeaux août 1974, pp. 66-123. IREM de Bordeaux.

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

[7] R. DUVAL, 1988, "Ecarts sémantiques et cohérence mathématique : introduction aux problèmes de congruence", *in Annales de Didactique et et Sciences Cognitives,* 1, IREM de Strasbourg.

[8] 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 26e rencontre, Bordeaux août 1974, 2-25. IREM de Bordeaux.

[9b] R.R. FALK, 1989, "Inference under uncertainty via conditional probabilities", *Studies in mathematics education. The teaching of statistics*, Vol. 7, pp. 175-184, Unesco, Paris.

[9b] R.R. FALK - I. LEVIN, 1980, "A potentiel for learning probability in young children", *Educational Studies in Mathematics*, Vol. 11, n°2, 181-204.

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

[11] B.V. GNEDENKO & A. KINTCHINE, 1960, *Introduction à la théorie des probabilités,* Monographie Dunod.

[12] 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.

[13] 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.

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

[15] M. et A. HENRY, 1992, "L'Enseignement des probabilités dans le programme de première", *Repères* n° 6, 27-52.

[16] M. HENRY, 1993, "Introduction à la théorie des probabilités et à son enseignement", *Repères* n°13., et Actes des colloques inter-IREM Commission statistique.

[17] A. KOLMOGOROV, 1933, *Foundations of the theory of tprobability,* Tchesla publishing company, New York(1956).

[18] 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.

[19] A. MESQUITA, 1989, *L'influence des aspects figuratifs dans l'argumentation des élèves en géométrie : élément pour une typologie*, Thèse de Doctorat, ULP Strasbourg I.

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

[21] PARZYSZ, 1993-a, *Notation et sens : une étude de cas,* Brochure de l'IREM de Lorraine.

[22] PARZYSZ, 1993-b, "Des statistiques aux probabilités : exploitons les arbres", *Repères* n° 10, 91-104.

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

[24a] J.M. SHAUGHENESSY, 1992, "Research in probability and statistics : reflections and directions", *Handbook of research on mathematics teaching and learning,* N.C.T.M. Douglas A.Grouws, Editor, pp. 465-494.

[24b] 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, 295-316.

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

[26] A. TOTOHASINA, 1992, "Conceptions causaliste ou chronologiste de la notion de probabilité conditionnelle", in *Cahiers de Didactique des Mathématiques*, fasc.9, Département de Mathématique de l'Institut Français de Thessalonique.

[27] 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.

[29] 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), 33-35, IREM de Strasbourg.

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

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

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

[33] M. ZAKI, 1990, *Traitements de problèmes de probabilités en situation de simulation,* Thèse de Doctorat d'université, ULP Strasbourg I.