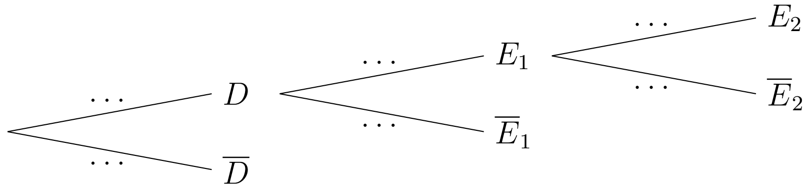
Arbres probabilistes : outil de résolution ou objet d'enseignement indispensable ?

Avant tout développement, quelques précisions de vocabulaire semblent nécessaires.
D'une part il s'agit de bien distinguer les trois types d'arbres mentionnés dans les programmes, documents ressource et manuels scolaires concernant les probabilités : *arbres des issues possibles*, *arbres de dénombrement* dont les branches sont pondérées par le nombre de façons d'obtenir ces issues et *arbres probabilistes* dont les branches sont pondérées par des probabilités.
D'autre part, dans la suite nous serons amenés à utiliser un concept qui nous semble fondamental pour l'enseignement des probabilités discrètes au collège et au lycée. Il s'agit de l'urne de Bernoulli [1] qui est une urne idéale composée de N boules telle que la probabilité de tirer l'une quelconque de ces boules est $1/N$. Dans une telle urne, la probabilité de tirer une boule d'une certaine catégorie est alors la proportion de boules de cette catégorie. Elle est archétypale du modèle d'équiprobabilité (trop souvent sous-entendu dans les manuels par la formule « au hasard ») et elle permet de modéliser des expériences aléatoires d'univers fini [2] : la réalisation d'un événement de probabilité k/N correspond au tirage d'une boule blanche dans une urne de Bernoulli contenant N boules dont k blanches.

Au moment où nous rédigeons cet article, les références aux arbres dans l'enseignement des probabilités se multipliant, il nous est apparu important de nous interroger sur leur statut dans les connaissances attendues d'un élève de collège ou de lycée et sur la nécessité de leur utilisation dans le cursus du lycée en nous appuyant sur la lecture de documents officiels.

D'où la question posée en titre de l'article qui peut sembler surprenante, voire saugrenue, tant est répandu l'usage d'arbres probabilistes tout d'abord dans les manuels, puis la préconisation de leur étude dans les deux (ou trois) dernières moutures des programmes.
Ainsi voit-on apparaître dans de nombreux exercices au baccalauréat des questions spécifiques à propos de la création d'arbres probabilistes qui deviennent un passage obligé pour la résolution de problème. Un exemple parmi tant d'autres [3] :

> Pour embaucher ses cadres une entreprise fait appel à un cabinet de recrutement. La procédure retenue est la suivante. Le cabinet effectue une première sélection de candidats sur dossier. 40 % des dossiers reçus sont validés et transmis à l'entreprise. Les candidats ainsi sélectionnés passent un premier entretien à l'issue duquel 70 % d'entre eux sont retenus. Ces derniers sont convoqués à un ultime entretien avec le directeur des ressources humaines qui recrutera 25 % des candidats rencontrés.
>
> **1.** On choisit au hasard le dossier d'un candidat.
> On considère les évènements suivants :
> — D : « Le candidat est retenu sur dossier »,
> — E_1 : « Le candidat est retenu à l'issue du premier entretien »,
> — E_2 : « Le candidat est recruté ».
>
> **(a)** Reproduire et compléter l'arbre pondéré ci-dessous.



---

[1] Nommée ainsi car Jacques Bernoulli l'utilise pour démontrer sa loi des grands nombres. Voir l'article *A propos de la définition en probabilité* de Jean-Claude Thiénard dans *Enseigner les probabilités au lycée*[4]

[2] Dans le cas de probabilités rationnelles.

[3] Sujet Bac S, Métropole 21 juin 2012, extrait de l'exercice 2.

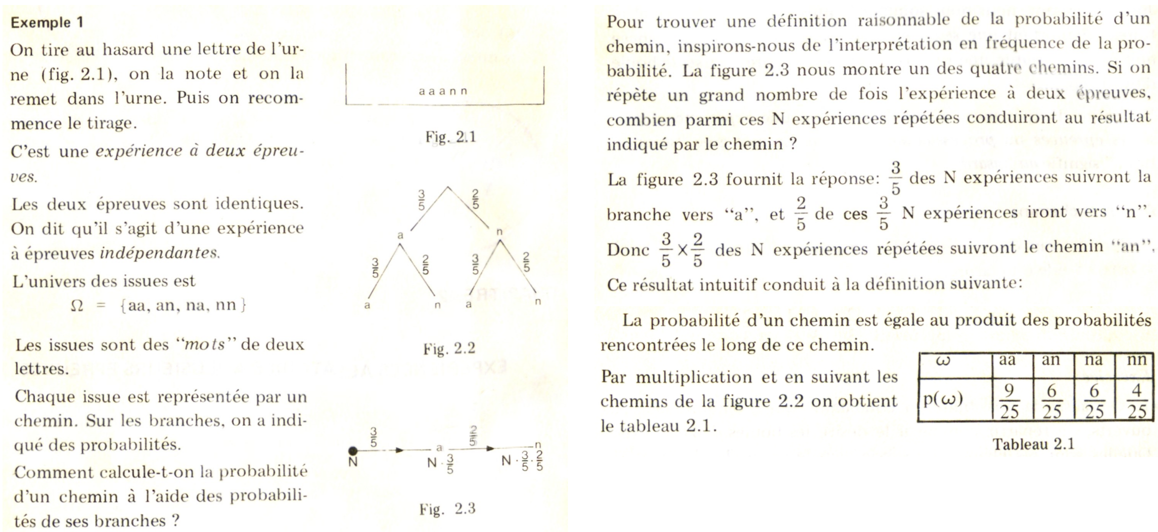
Notre interrogation peut se décliner en diverses questions auxquelles nous allons tenter d'apporter des réponses :

- l'arbre probabiliste est-il un outil incontournable pour résoudre de nombreuses situations ?
- comment peut-on envisager l'enseignement des arbres probabilistes ?
- l'enseignement de la notion d'arbre probabiliste ne peut-il pas aboutir à la création d'obstacles didactiques ?
- comment en est-on arrivé à faire figurer l'étude des arbres probabilistes dans les programmes ?
- les arbres probabilistes apparaissent-ils dans le savoir académique [4] ?

La dernière question que nous nous posons est vite réglée : dans tous les écrits fondateurs de la théorie des probabilités que nous avons consultés, il n'est jamais question d'arbres probabilistes [5]. Cela tendrait à prouver que les arbres probabilistes sont une création didactique.

I. Quand les arbres probabilistes sont-ils apparus dans la littérature pédagogique ?

Engel en donne une des premières formalisations à l'usage des enseignants dans son ouvrage *L'enseignement des probabilités et statistiques* [6] (page 36) après qu'il ait introduit les arbres de dénombrement (page 20).
Sa « justification » des calculs basés sur les arbres probabilistes repose sur une approche fréquentiste de la probabilité, sur la notion d'indépendance (non définie mais nous en reparlerons ultérieurement) et sur le calcul de proportions de proportions.



Exemple 1

On tire au hasard une lettre de l'urne (fig. 2.1), on la note et on la remet dans l'urne. Puis on recommence le tirage.

C'est une *expérience à deux épreuves*.

Les deux épreuves sont identiques. On dit qu'il s'agit d'une expérience à épreuves *indépendantes*.

L'univers des issues est

$$\Omega = \{aa, an, na, nn\}$$

Les issues sont des "*mots*" de deux lettres.

Chaque issue est représentée par un chemin. Sur les branches, on a indiqué des probabilités.

Comment calcule-t-on la probabilité d'un chemin à l'aide des probabilités de ses branches ?

Fig. 2.1

Fig. 2.2

Fig. 2.3

Pour trouver une définition raisonnable de la probabilité d'un chemin, inspirons-nous de l'interprétation en fréquence de la probabilité. La figure 2.3 nous montre un des quatre chemins. Si on répète un grand nombre de fois l'expérience à deux épreuves, combien parmi ces N expériences répétées conduiront au résultat indiqué par le chemin ?

La figure 2.3 fournit la réponse: $\frac{3}{5}$ des N expériences suivront la branche vers "a", et $\frac{2}{5}$ de ces $\frac{3}{5}$ N expériences iront vers "n". Donc $\frac{3}{5} \times \frac{2}{5}$ des N expériences répétées suivront le chemin "an".

Ce résultat intuitif conduit à la définition suivante:

La probabilité d'un chemin est égale au produit des probabilités rencontrées le long de ce chemin.

Par multiplication et en suivant les chemins de la figure 2.2 on obtient le tableau 2.1.

ω	aa	an	na	nn
$p(\omega)$	$\frac{9}{25}$	$\frac{6}{25}$	$\frac{6}{25}$	$\frac{4}{25}$

Tableau 2.1

[4] Ce savoir est nommé « savoir savant » dans les écrits de Yves Chevallard.

[5] Huygens utilise une présentation sous forme d'arbre pour résoudre son 5$^{\text{ème}}$ exercice voir *Epistémologie et Histoire des Mathématiques*[2] mais d'une part ce n'est pas un arbre probabiliste au sens défini ci-avant et d'autre part les auteurs suivants, comme Bernoulli, qui reprendront son ouvrage, ne conserveront pas cette présentation.

[6] *L'enseignement des probabilités et statistiques*[3]

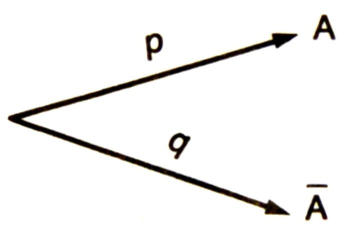
On trouve trace d'arbres probabilistes dans certains manuels des années 1980 mais de manière anecdotique. Il en est tout autrement des arbres de dénombrement car jusque dans les années 2000, l'enseignement des probabilités était fondé sur la combinatoire.

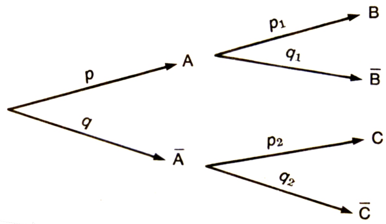
II. Quand les arbres sont-ils apparus dans les programmes ?

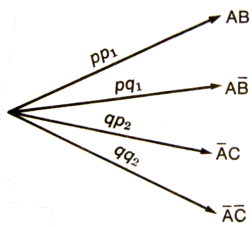
Nous avons regardé les programmes depuis 1974 en notant les évolutions significatives.

En 1974, les programmes des séries scientifiques restent ceux de la décennie précédente. Les probabilités y sont abordées de manière théorique. Dans les programmes figure alors aussi l'étude des ensembles, des cardinaux, des arrangements, permutations, injections...
Il n'est pas question de représentation (en particulier par les arbres) ni dans les programmes, ni dans les instructions.

En 1982, la partie probabilité est fortement marquée par l'étude de la combinatoire. Cependant pour la première fois dans l'annexe au commentaire du programme de Terminale B des approches possibles du schéma de Bernoulli sont présentées :
- la première : constitution de l'urne équivalente à deux tirages ;
- la seconde : **introduction par des arbres :** deux branches, puis deux fois deux branches que l'on regroupe en un arbre, puis arbre à quatre branches équivalent, enfin schéma de Bernoulli et généralisation à n répétitions.

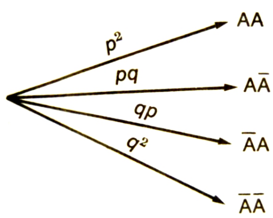




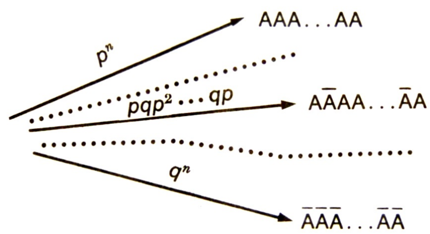


« Les probabilités portées par les branches étant le produit des probabilités correspondantes ».

Dans le cas du Schéma de Bernoulli on obtient :



puis



En revanche, rien n'est indiqué dans les commentaires du programme de terminale D, qui lui comporte des probabilités conditionnelles.
Les arbres probabilistes sont indiqués comme un outil possible de résolution.

En 1991, les probabilités conditionnelles et le schéma de Bernoulli apparaissent dans les programmes de terminales C et E, ainsi que la notion de variable aléatoire. Les premières notions de probabilités (réunion, intersection, contraire) sont maintenant introduites en première. Les « modèles d'urne » restent présents dans les exemples de TP. Mais surtout on voit arriver explicitement le terme d'arbre : « exemples simples d'emploi de partitions et de représentation (arbres, tableaux, ...) pour organiser et dénombrer des données ».
Notons que ce sont des arbres de dénombrement et non des arbres probabilistes.

En 1997, l'usage des arbres se généralise, on le trouve régulièrement dans la marge de droite en commentaire des contenus : « on exploitera la représentation en arbre ». Ils sont indiqués non seulement en combinatoire pour dénombrer (comme dans le programme précédent), mais surtout les arbres pondérés apparaissent pour le calcul des probabilités : « on introduira les arbres pondérés. Prenant appui sur le travail fait en première, on en explicitera les règles de fonctionnement pour leur utilisation comme outil de calcul ». Dans ce programme, on voit aussi apparaitre la notation $p_B(A)$ pour les probabilités conditionnelles. En revanche, la formule des probabilités totales n'apparait plus dans les contenus, elle est signalée en marge dans le programme de terminale ES mais pas dans celui de S ; elle semble être devenue un implicite des calculs de probabilités... peut-être découlant naturellement de l'exploitation des règles de calculs sur les arbres pondérés ?
La tendance à exploiter les arbres probabilistes se généralise.

En 2001, les arbres restent indiqués comme des représentations utiles au calcul des probabilités. Ils deviennent plus centraux dans le raisonnement : « un arbre de probabilité correctement construit constitue une preuve ».
Le formalisme des mathématiques dites modernes a totalement disparu. Les arbres probabilistes qui étaient un moyen didactique pour enseigner les probabilités acquièrent un autre statut.

En 2009 (et suivantes), le dénombrement a disparu des programmes. Les arbres apparaissent dès la classe de seconde, sans mention de pondération ; on peut donc penser qu'il s'agit d'arbres de dénombrement ou d'arbres des issues possibles. En première, « l'arbre pondéré est la représentation *privilégiée* pour représenter la répétition d'expériences identiques et indépendantes », il s'agirait donc d'un arbre de probabilité. Enfin en terminale, l'arbre pondéré est une capacité attendue et comme en 2001, « un arbre pondéré correctement construit constitue une preuve ».

Ainsi, l'étude des programmes depuis une quarantaine d'années montre comment les arbres probabilistes, conçus d'abord comme un moyen parmi d'autres pour enseigner les probabilités, sont devenus un objet d'enseignement incontournable ne serait-ce que pour répondre aux questions posées au baccalauréat. Ils sont d'abord apparus (si on omet les arbres de dénombrement) pour calculer des probabilités en cas d'épreuves répétées (en particulier dans le schéma de Bernoulli) puis dans les calculs utilisant des probabilités conditionnelles. À regarder les manuels actuels, il n'y aurait même que les arbres pour enseigner les probabilités conditionnelles [7]. Or il en est tout autrement.

III. Comment enseigner actuellement les arbres probabilistes en suivant les préconisations des documents ressources ?

1. Au collège

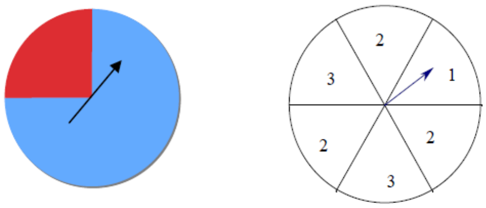
En 2008, les probabilités apparaissent dans les programmes du collège :
« La notion de probabilité est abordée à partir d'expérimentations qui permettent d'observer les fréquences des issues dans des situations familières (pièces de monnaie, dés, roues de loteries, urnes, etc). La notion de probabilité est utilisée pour modéliser des situations simples de la vie courante. Les situations étudiées concernent les expériences aléatoires à une ou à deux

[7] Nous conseillons de lire l'article sur le tableau de signes, APMEP, bulletin vert http://www.apmep.fr/IMG/pdf/bull-474_Gaud.pdf pour voir le parallélisme entre l'évolution de la notion d'arbre probabiliste et celle du tableau de signes d'une expression polynomiale.

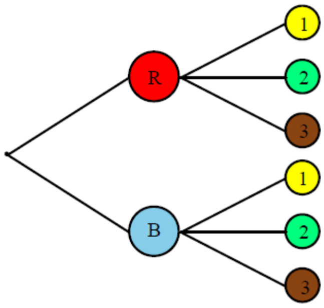
épreuves. »

Aucune référence aux arbres de quelques natures que ce soit dans l'énoncé du programme mais, la question de savoir comment enseigner les expériences aléatoires à deux épreuves se posant, il en est tout autrement dans le document d'accompagnement dont voici un extrait [8] :

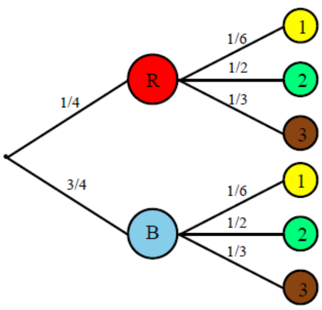
> Aucune compétence sur les expériences à deux épreuves n'est exigible dans le cadre du socle commun. Les situations proposées aux élèves doivent rester élémentaires. On se bornera à des expériences conduisant à un maximum de 6 cas (2×2, 2×3 ou 3×2) et on n'abordera pas les cas de tirages successifs dans une urne (avec ou sans remise)[9].
>
> On considère l'expérience suivante, qui se déroule en deux étapes : d'abord, on fait tourner la roue de loterie située ci-dessous à gauche (on obtient la couleur « Rouge » avec une probabilité de 0,25 et la couleur « Bleu » avec une probabilité de 0,75). Ensuite, on fait tourner la deuxième roue de loterie (on obtient le numéro 1 avec la probabilité 1/6, le numéro 2 avec la probabilité 1/2 et le numéro 3 avec la probabilité 1/3).



> Un arbre non pondéré permet de déterminer tous les résultats possibles à l'issue de ces deux étapes. Les résultats possibles peuvent être notés ainsi : (R, 1), (R, 2), (R, 3), (B, 1), (B, 2), (B, 3). Chacun de ces résultats est représenté dans l'arbre ci-dessous par la succession de deux branches[10].



> On cherche la probabilité d'obtenir chacun des six résultats possibles à l'issue des deux étapes : (R, 1), …, (B, 3).
> On peut pondérer l'arbre avec les probabilités :



La démarche rejoint celle de Engel et suscite quelques remarques et questions :

- l'exemple choisi entretient la confusion entre le modèle et la réalité ;
- l'approche fréquentiste est mathématiquement problématique : la probabilité est un nombre qui est probablement voisin de la fréquence. La probabilité est définie par les probabilités : on a donc une définition circulaire ;
- l'approche fréquentiste ne serait-elle pas incompatible avec la notion d'intervalle de fluctuation enseigné en seconde ?

[8] http://www.arpeme.fr/documents/7E7DAF5DD2B58FE9343.pdf

— la justification par la multiplication des fréquences est plus de nature psychologique que logique ;
— l'apprentissage dès la troisième de la formule $P(A \cap B) = P(A) \times P(B)$ n'est-il pas source d'obstacle dans les classes ultérieures lorsque sera introduite la notion de probabilités conditionnelles ?
— la notion d'indépendance est laissée à l'appréciation du lecteur.

2. Au lycée

Voici ce que disent les documents ressources[9] :

A. Exemple d'expérience aléatoire à deux épreuves

On se donne :

- *une urne contenant quatre boules indistinguables au toucher dont trois boules bleues, notées b1, b2 et b3, portant respectivement les numéros 1, 2 et 3, et une boule rouge unique, notée r ;*
- *un jeu de six cartes identiques portant chacune un chiffre en couleur : une carte avec un chiffre "1" en vert, une carte avec un chiffre "2" en rouge, une carte avec un chiffre "2" en bleu, une carte avec un chiffre "2" en vert, une carte avec un chiffre "3" en rouge, une carte avec un chiffre "3" en bleu.*

On considère l'expérience aléatoire suivante : on prélève de façon équiprobable une boule dans l'urne puis une carte du jeu. On note, dans l'ordre, la couleur de la boule extraite et le numéro inscrit sur la carte. On rappelle qu'un modèle associé à cette expérience aléatoire est défini par la donnée :

- *de l'ensemble Ω de toutes les issues possibles de l'expérience ;*
- *d'une probabilité P déterminée par ses valeurs pour chacun des événements élémentaires définis par ces issues.*

La liste de toutes les issues possibles peut être trouvée en utilisant l'arbre des possibles ci-dessous. Les issues possibles pour cette expérience aléatoires sont les couples (R,1) ; (R,2) ; (R,3) ; (B,1) ; (B,2) ; (B,3) où B désigne la couleur « Bleu » et R la couleur « Rouge ».

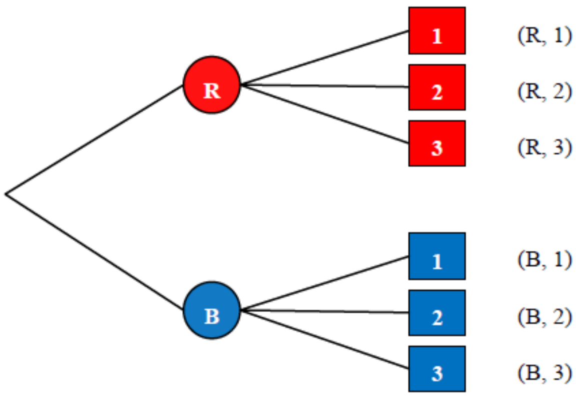


Figure 1

Une fois les issues toutes identifiées, il s'agit de trouver la probabilité des événements élémentaires déterminés par chacune des issues. Il est clair que l'équiprobabilité n'est pas une réponse possible. En effet, on a des raisons de penser que la couleur « Bleu » sera plus probable que la couleur « Rouge » et que le chiffre "2" a plus de chances de

[9] http://cache.media.eduscol.education.fr/file/Mathematiques/59/6/Ressource_Statistiques_Probabilites_1eres_208596.pdf

sortir que les autres ; en conséquence, l'issue (B,2) a plus de chances de sortir que l'issue (R,1).
Pour affecter une probabilité à chacune des issues, nous allons considérer un autre modèle (qualifié par la suite de ***modèle intermédiaire****) qui prend en compte, pour la boule extraite, sa couleur et aussi son numéro éventuel, et pour la carte, le chiffre mentionné mais aussi sa couleur. On peut recenser tous les résultats par l'arbre représentant les issues possibles ci-après.*

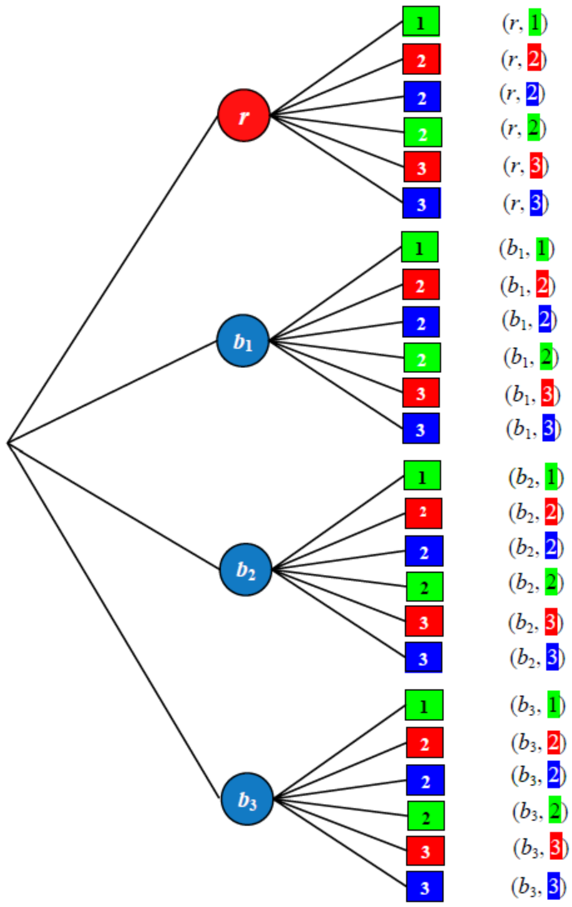


Figure 2

On obtient 4×6 *résultats possibles. On peut les noter de la façon suivante :* $(r, 1)$ *;* $(r, 2)$ *;* $(r, 2)$ *;* $(r, 2)$ *;* $(r, 3)$ *;* $(r, 2)$ *; [. . .]*
Chaque branche de l'arbre représente une issue, et compte tenu des conditions du tirage équiprobable de la boule, puis du tirage équiprobable de la carte, il n'y a pas de raison de penser qu'une branche de l'arbre ait plus de chances d'être parcourue qu'une autre. On peut donc considérer que chacune des issues précédentes a la même probabilité, égale à $1/24$*, d'être réalisée. Dans le modèle intermédiaire, par exemple, l'événement « Tirer une boule bleue puis une carte portant le chiffre "2" » se représente mathématiquement par le sous-ensemble des issues* $\{(b_1, 2); (b_1, 2); (b_1, 2); (b_2, 2); (b_2, 2); (b_2, 2); (b_3, 2); (b_3, 2); (b_3, 2)\}$.
Par suite, la probabilité de cet événement sera égale à $9/24$*. Revenant alors au premier modèle où l'événement « Tirer une boule bleue puis une carte portant le chiffre "2" » se représente mathématiquement par l'événement élémentaire {(B,2)}, on prendra* $9/24$ *pour la probabilité d'obtenir l'issue (B,2). On peut faire de même pour les cinq autres issues : (R,1) ; (R,3) ; (B,1) ; (B,2) ; (B,3).*

La démarche menée avec les arbres de dénombrement permet de trouver le modèle de l'urne de Bernoulli donc de calculer les probabilités sans passer par les arbres probabilistes.

B. Justification de l'arbre des probabilités

Si on revient à l'arbre (cf. figure 2) utilisé pour trouver toutes les issues possibles du modèle intermédiaire, on constate que cet arbre est très fastidieux à dessiner. Dans la mesure où on ne s'intéresse qu'à la couleur de la boule et au chiffre inscrit sur la carte, on peut alléger sa construction, moyennant quelques conventions de lecture, pour retrouver l'arbre (cf. figure 1) des issues possibles du premier modèle pondéré par les probabilités et justifier la règle des produits de la façon suivante :

Étape 1 :

Partant de l'arbre de la figure 2, dans la mesure où on ne s'intéresse qu'à la couleur de la boule (et non à son numéro éventuel) et qu'au chiffre inscrit sur la carte (et non à sa couleur) on peut convenir de représenter chaque branche de l'arbre de la figure 2 aboutissant à la même couleur de boule, par une seule branche comprenant autant de traits parallèles qu'il y a de boules physiques de cette même couleur. On procède de même en représentant chaque branche de l'arbre de la figure 2 aboutissant à un même chiffre de carte, par une seule branche comprenant autant de traits parallèles qu'il y a de cartes physiques avec ce même chiffre inscrit avec des couleurs différentes. On obtient ainsi l'arbre plus simple de la figure 3 ci-après qui contient cependant autant de branches que celui de la figure 2 tout se rapprochant de l'allure de l'arbre de la figure 1.

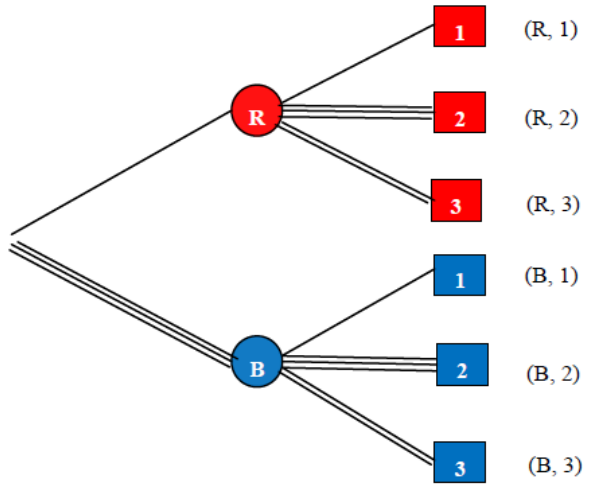


Figure 3

Étape 2

On peut alors simplifier davantage l'arbre de la figure 3, en représentant chaque branche par un seul trait pondéré par le nombre de traits composant la branche correspondante dans l'arbre de la figure 3. On obtient alors l'arbre pondéré de la figure 4 qui suit :

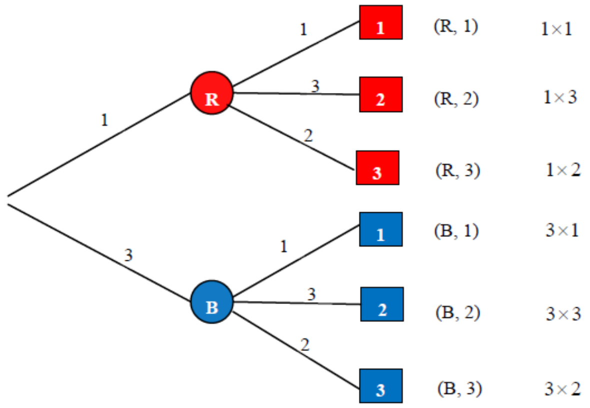


Figure 4

On remarque alors que le produit des nombres rencontrés le long d'un chemin représentant une issue du premier modèle est égal au nombre de chemins de l'arbre de la figure 2 qui réalisent l'événement correspondant dans le modèle intermédiaire. Ainsi, pour l'événement « Tirer une boule bleue puis une carte portant le chiffre 2 », c'est-à-dire (B,2), le produit 3×3 *est égal au nombre de chemins dans le modèle intermédiaire, soit 9.*

On a simplifié l'arbre de dénombrement qui, comme cela était déjà le cas avec l'arbre de la figure 2, donne la composition de l'urne de Bernoulli permettant de calculer la probabilité de chacune des issues : la probabilité de l'issue (B,2) est 9/24.

Comment le document ressources invite-t-il à passer de l'arbre de dénombrement à l'arbre probabiliste ?

Étape 3
Cette étape consiste à pondérer chaque branche de l'arbre, non plus avec le nombre de traits composant la branche correspondante dans l'arbre de la figure 3, mais avec le quotient de ce nombre par le nombre total de branches d'un même niveau. On obtient ainsi l'arbre pondéré suivant :

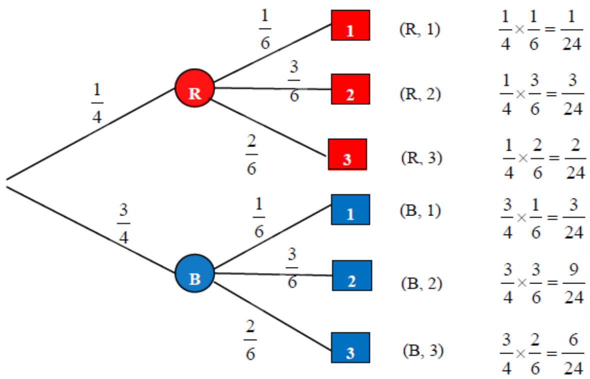


Figure 5

<u>On remarque</u>[10] *alors que le produit des quotients affectés aux diverses branches d'un chemin aboutissant à une issue donnée du premier modèle, par exemple pour (B,2) le produit* $3/6 \times 3/4$ *, est égal à la probabilité, dans cet exemple* $9/24$*, que cette issue se réalise. Cette remarque est valable pour toutes les branches de l'arbre. Au final, on peut constater que l'arbre de la figure 5 n'est rien d'autre que l'arbre de probabilités associé à l'arbre des possibles de la figure 1.*

L'étape 3, passage de l'arbre de la figure 4 à celui de la figure 5, est problématique car une fois encore fondée sur un argument plus psychologique que logique : « on remarque que ». Elle apparaît comme un tour de passe-passe pédagogique qui permet la justification de la multiplication des probabilités sans utiliser les probabilités conditionnelles. Mais cela fonctionne-t-il dans toutes les situations ?
La justification est complètement différente de celle proposée en troisième basée il est vrai sur l'approche fréquentiste préconisée par les programmes. Comment fait-on cohabiter ces deux approches ?
D'autre part, au lycée, il semble, sans que cela ne soit explicité, que les calculs de probabilités sont fondés sur le modèle de l'urne de Bernoulli. Ce modèle, que l'on perd de vue dans la démarche proposée par le document ressource, permet de calculer immédiatement les probabilités cherchées à partir de l'arbre de la figure 4 sans utilisation d'arbre probabiliste.

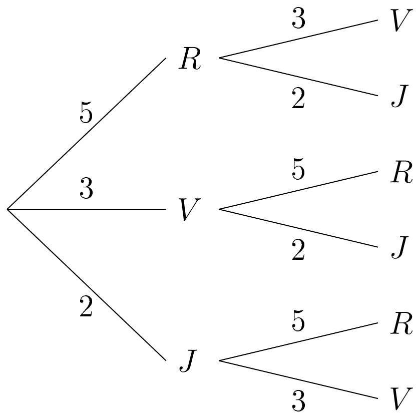
[10] Souligné par les auteurs.

IV. Le doute

Considérons l'exercice suivant :
Une urne est composée de cinq boules rouges, trois boules vertes et deux boules jaunes invisibles de l'extérieur et indiscernables au toucher (ce qui assure l'équiprobabilité des tirages d'une boule dans cette urne). L'épreuve consiste à tirer une boule, noter sa couleur, enlever de l'urne toutes les boules de cette couleur puis en tirer une deuxième et noter sa couleur.
Calculer la probabilité de tirer une boule jaune puis une verte.

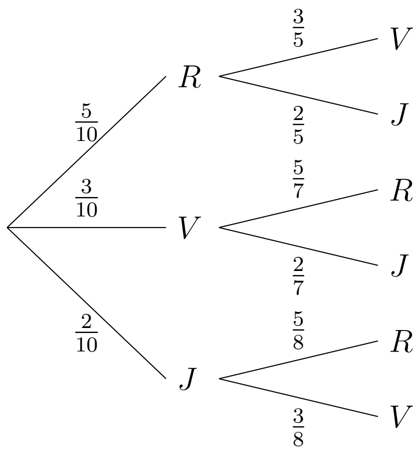
Pour résoudre cet exercice en classe de première, reprenons le document d'accompagnement :

Dresser l'arbre des issues possibles conformément au document ressource, puis pondérer chaque branche par le nombre de boules de chaque couleur selon la composition de l'urne dans laquelle s'effectue le tirage :



Arbre 1 qui correspond à l'arbre de la figure 4 du document ressource

Puis passer à l'arbre probabiliste en suivant l'étape 3 du document ressource :



Arbre 2 qui correspond à l'arbre de la figure 5 du document ressource

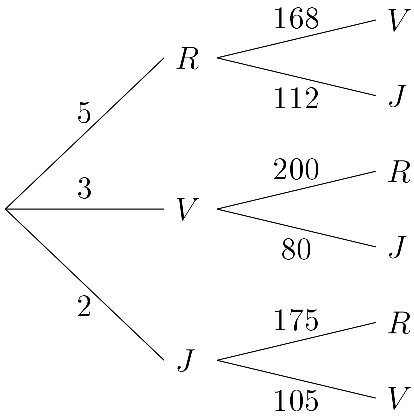
La probabilité de tirer une boule jaune puis une verte est donc $2/10 \times 3/8 = 3/40$.

Mais qu'en est-il de la justification donnée dans le document ressource ?

> Pour rappel : « ***On remarque*** *alors que le produit des quotients affectés aux diverses branches d'un chemin aboutissant à une issue donnée du premier modèle, par exemple pour (B,2) le produit* $3/6 \times 3/4$*, est égal à la probabilité, dans cet exemple* $9/24$*, que cette issue se réalise* ».

D'après l'arbre 1 correspondant à celui de la figure 4 du document ressource, on obtient 15 RV, 10 RJ, 15 VR, 6 VJ, 10 JR et 6 JV. D'où on pourrait conclure que sur 62 issues il y en a 6 JV donc la probabilité de tirer une boule jaune puis une verte est 6/62 soit 3/31.
On observe alors que la justification donnée dans le document ressource ne fonctionne plus (tout en sachant que le résultat 3/40 qui repose en fait sur les probabilités conditionnelles est exact).

Pour comprendre l'erreur commise dans le raisonnement lié à l'arbre 1, il faut remarquer que l'urne constituée suite au dénombrement, {15 RV, 10 RJ, 15 VR, 6 VJ, 10 JR et 6 JV}, n'est pas une urne de Bernoulli dans la mesure où les issues recensées ne sont pas équiprobables. En effet la proportion d'issues avec R en premier (15/62) n'est pas égale à la probabilité de tirer une boule rouge au premier tirage (1/2), de même pour les issues avec V en premier et avec J en premier.
Néanmoins on peut envisager un raisonnement utilisant une urne de Bernoulli modélisant l'épreuve à condition de constituer cette urne en respectant les probabilités de tirages avec R, V ou J en premier. Pour cela il suffit que le nombre total de boules de l'urne dans laquelle on effectue le second tirage soit le même, quel que soit l'issue du premier tirage, par exemple le ppcm de 5, 7 et 8 soit 280. On peut alors dresser l'arbre de dénombrement en respectant la condition énoncée ci-dessus et en affectant aux branches du second niveau une pondération conforme aux probabilités du second tirage :



Arbre 3

Sur cet arbre, on dénombre 1400 issues avec R en premier, 840 avec V en premier et 560 avec J en premier sur un total de 2800. Les proportions correspondent alors aux probabilités d'obtention de chaque couleur au premier tirage.
D'où à l'aide de l'urne de Bernoulli qui en découle, {840 RV, 560 RJ, 600 VR, 240 VJ, 350 JR, 210 JV}, on obtient pour probabilité de JV : 210/2800 soit 3/40.

Autrement dit la justification s'appliquant à l'exemple choisi dans le document ressource s'avère fausse.

V. Comment enseigner le calcul de probabilités en cas d'épreuve successives [11]

Il nous semble que, quel que soit le niveau où les probabilités sont enseignées, il ne faut pas confondre les expériences réelles et le modèle et que le calcul ne peut s'effectuer qu'une fois un

[11] L'intérêt des arbres probabilistes réside dans le calcul de probabilités dans le cas d'épreuves successives.

modèle choisi. L'approche fréquentiste ne donne pas un modèle, il permet d'en choisir un. Seule l'urne de Bernoulli permet d'effectuer des calculs.
Jusqu'à l'introduction des probabilités conditionnelles, le calcul de probabilités lors d'épreuves successives ne peut passer que par la construction de l'urne de Bernoulli en s'aidant éventuellement d'arbres de dénombrement. Autrement dit nous pensons qu'il faut éviter les arbres probabilistes avant la terminale.
La multiplication des probabilités en cas d'épreuves successives ne peut être comprise que par le truchement des probabilités conditionnelles introduites à l'aide des urnes. L'utilisation prématurée des arbres probabilistes conduisant à la formule $P(A \cap B) = P(A) \times P(B)$ risque d'engendrer des automatismes préjudiciables en terminale quand il n'y a plus indépendance des épreuves. On peut à ce propos être surpris de trouver dans les manuels ce qui suit sans aucune justification[12] :

1.2 Cas général

Lorsqu'une expérience aléatoire est la répétition de n épreuves identiques et indépendantes, on peut la représenter par un arbre pondéré.
On admettra que :
- une issue est une liste ordonnée de résultats, représentée par un chemin ;
- la probabilité d'une issue est le produit des probabilités de chacun de ces résultats.

1.3 Règles d'utilisation d'un arbre pondéré

Les règles suivantes précisent les conditions à respecter pour construire un arbre pondéré représentant une expérience aléatoire et calculer les probabilités des événements.

Règle 1 La somme des probabilités inscrites sur les branches issues d'un même nœud vaut 1.
Ceci est la loi des nœuds.

Règle 2 La probabilité d'une issue représentée par un chemin est égale au produit des probabilités inscrites sur les branches de ce chemin.
Ceci est la loi des chemins.

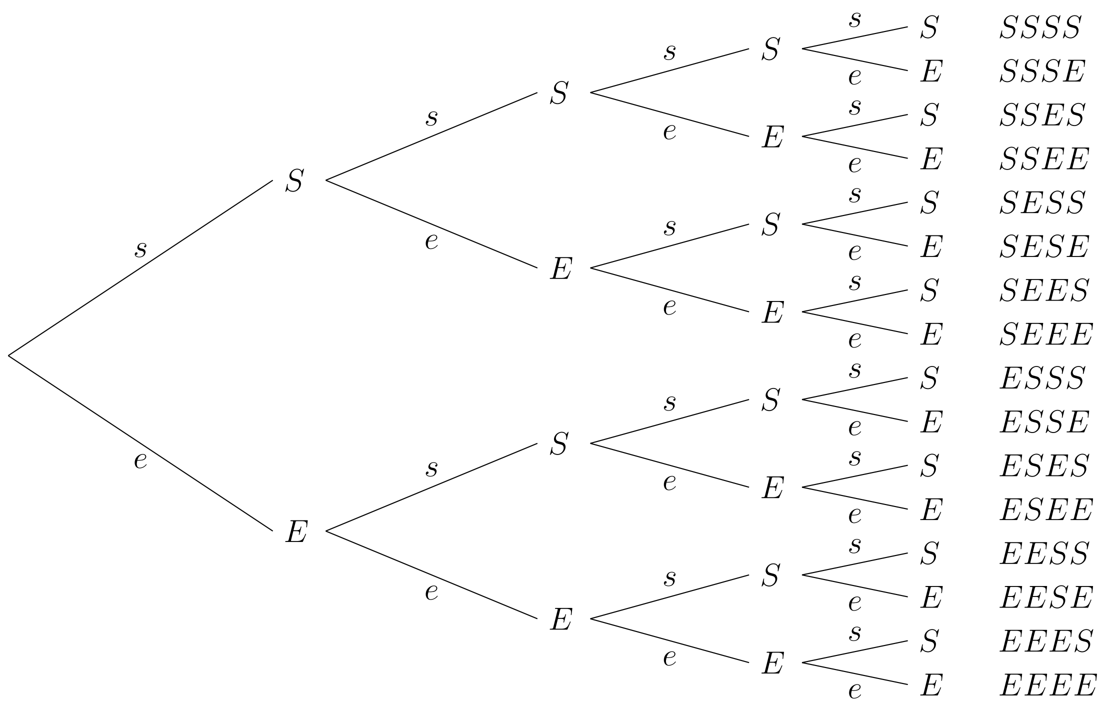
Règle 3 La probabilité d'un événement A est la somme des probabilités des issues associées aux chemins qui conduisent à la réalisation de A.

En classe de première, jusqu'à l'introduction de la loi binomiale, le modèle de l'urne de Bernoulli permet de traiter les exercices à propos d'épreuves successives.
En ce qui concerne la loi binomiale, les arbres probabilistes, comme nous venons de le voir n'étant pas justifiés de manière satisfaisante sans la notion de probabilités conditionnelles, nous proposons, ci-dessous, deux démarches possibles :

— **Première démarche :** Découvrir la loi binomiale sans les arbres **probabilistes**
Soit une épreuve $\mathcal{E}$ à deux issues, l'une appelée succès notée S de probabilité p et l'autre échec notée E de probabilité $1-p$. $\mathcal{E}$ peut être modélisée par une urne de Bernoulli composée de N boules, s marquées S et $e = N - s$ boules marquées E avec $p = \frac{s}{N}$ et $1 - p = \frac{e}{N}$.
Considérons l'expérience aléatoire $\mathcal{M}$ consistant à répéter quatre fois de suite l'épreuve $\mathcal{E}$, modélisée à chaque occurrence de $\mathcal{E}$ par l'urne de Bernoulli décrite ci-dessus[13].
On cherche à déterminer la loi de probabilité de la variable aléatoire X égale au nombre de succès obtenus lors de la réalisation de l'expérience $\mathcal{M}$. Il s'agit alors de déterminer une urne de Bernoulli modélisant $\mathcal{M}$, ce qui peut se faire à l'aide d'un arbre de dénom-

[12] Transmath 1ereS, programme 2011, page 320[1]
[13] Cette formulation est une manière de comprendre la notion de répétition d'épreuves identiques et indépendantes figurant dans le programme.

brement comme celui de la figure 4 du document ressource, basé sur le principe de la multiplication.



On obtient ainsi une urne de Bernoulli[14] modélisant $\mathcal{M}$, comportant N^4 boules, s^4 marquées $SSSS$, s^3e marquées $SSSE$, s^2e^2 marquées $SSEE$, ...
Sur l'arbre, un seul chemin mène à quatre succès mais on remarque que plusieurs chemins mènent à des issues comportant le même nombre de succès (et d'échecs) c'est-à-dire à une même valeur de X.
Six chemins aboutissent à deux succès et deux échecs : chacun des chemins correspond à s^2e^2 boules de l'urne de Bernoulli qui en compte donc $6 \times s^2e^2$ correspondant à $X = 2$.
Quatre chemins aboutissent à trois succès et un échec : chacun des chemins correspond à s^3e boules de l'urne de Bernoulli qui en compte donc $4 \times s^3e$ correspondant à $X = 3$...
En notant $\binom{4}{k}$ le nombre de chemins aboutissant à k succès lors de l'expérience $\mathcal{M}$, on a dans l'urne de Bernoulli modélisant $\mathcal{M}$, $\binom{4}{k} \times s^4$ boules correspondant à $X = 4$, $\binom{4}{3} \times s^3e$ boules correspondant à $X = 3$, $\binom{4}{2} \times s^2e^2$ boules correspondant à $X = 2$, ... D'où la probabilité que :

— $X = 4$ est $\binom{4}{4} \times \frac{s^4}{N^4} = \binom{4}{4} \times \left(\frac{s}{N}\right)^4 = \binom{4}{4} \times p^4 = p^4$

— $X = 3$ est $\binom{4}{3} \times \frac{s^3e}{N^4} = \binom{4}{3} \times \left(\frac{s}{N}\right)^3 \frac{e}{N} = 4 \times p^3(1-p)$

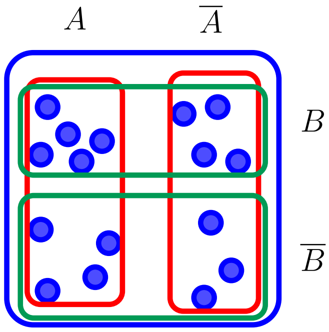
— $X = 2$ est $\binom{4}{2} \times \frac{s^2e^2}{N^4} = \binom{4}{2} \times \left(\frac{s}{N}\right)^2 \left(\frac{e}{N}\right)^2 = 6 \times p^2(1-p)^2$

...

Ainsi la loi de probabilité de X est, pour k entre 0 et 4, $P(X = k) = \binom{4}{k} p^k(1-p)^{4-k}$.
Formule qui peut être généralisée au cas de la répétition de n épreuves et ainsi obtenir la loi de probabilité d'une variable aléatoire suivant la loi n et p.

[14] On remarquera que l'arbre de dénombrement obtenu respecte bien les conditions permettant de définir l'urne de Bernoulli modélisant l'épreuve $\mathcal{M}$, contrairement à l'arbre 1 du paragraphe IV.

— **Deuxième démarche :** introduire les probabilités conditionnelles en première
Injonction est faite dans les programmes de ne pas introduire les probabilités conditionnelles mais il ne s'agit en aucun cas de traiter le programme de terminale, seulement justifier la multiplication des probabilités figurant sur les branches des arbres probabilistes. On peut procéder ainsi avec une urne de Bernoulli composée de N boules ayant deux caractères A et B.



Pour calculer la probabilité que B soit réalisé sachant que A est réalisé, on modélise la situation en se plaçant dans une nouvelle urne qui n'est composée que des boules marquées A et ainsi $P_A(B) = N_{A\cap B}/N_A = (N_{A\cap B}/N)/(N_A/N) = P(A \cap B)/P(A)$. Ainsi on retrouve la formule des probabilités conditionnelles.
Dans le cas de la loi binomiale, on doit alors admettre que $P_A(B) = P(B)$ car la réalisation de A ne peut pas avoir d'incidence sur réalisation de B.
On peut alors reprendre la démarche classique enseignée dans les manuels.

VI. Conclusion

Loin de nous l'idée de bannir l'usage des arbres probabilistes de notre enseignement. Notre propos a été de montrer qu'il est souhaitable de ne pas les introduire avant les probabilités conditionnelles au risque de créer des obstacles didactiques et qu'il est possible de le faire en n'utilisant que des arbres de dénombrement puis en calculant les probabilités après modélisation par une urne de Bernoulli. L'évolution du statut des arbres probabilistes dans les programmes montre qu'ils sont devenus objet d'enseignement alors qu'ils n'ont aucune vie dans le savoir académique. C'est ce que les didacticiens nomment glissement métacognitif. C'est malheureusement un phénomène classique[15] dans l'enseignement où un *outil* introduit dans les manuels devient un *objet d'étude* cité à étudier dans les programmes.
D'autre part, derrière les apparences se cache la notion d'indépendance. Elle est considérée comme naturelle depuis le collège sans être définie mathématiquement. Cela suffit largement pour résoudre les situations rencontrées dans les classes du second degré mais alors faut-il introduire l'indépendance de deux événements au lycée vu l'usage qui en est fait ?
Enfin, dernière remarque, le recours systématique aux arbres probabilistes peut installer chez certains élèves une conception chronologique ou causaliste de la notion de probabilités conditionnelles qui risque de constituer un obstacle didactique lorsqu'en terminale il sera demandé de « renverser » l'arbre[16].

[15] Par exemple, *Quelques interrogations à propos du « tableau de signes »* dans le bulletin vert APMEP n° 474, http://www.apmep.fr/IMG/pdf/bull-474_Gaud.pdf
[16] Voir André Totohasina, l'introduction du concept de probabilité conditionnelle : avantages et inconvénients de l'arborescence[8].

Références

[1] In : *Transmath 1re S*. Nathan, 2011. Chap. 13 : Loi binomiale, p. 320.

[2] Evelyne Barbin et Jean-Pierre Lamarche. In : *IREM - Epistémologie et Histoire des Mathématiques*. Ellipses, 2004, p. 25.

[3] Arthur Engel. *L'enseignement des probabilités et statistiques, volume 1*. Cedic, 1975.

[4] "Enseigner les probabilités au lycée". In : *Commission Inter-Irem « Statistiques et probabilités »* (1997).

[5] Bernard Parzysz. "Des statistiques aux probabilités, exploitons les arbres". In : *Repère IREM* (1993), p. 91–104.

[6] Bernard Parzysz. "Un outil sous estimé : l'arbre probabiliste". In : *bulletin APMEP no 372* (1990), p. 47–52.

[7] Heinz Steinbring. "L'indépendance stochastique". In : *Recherche en didactique des mathématiques 7.3* (1986).

[8] André Totohasina. "L'introduction du concept de probabilité conditionnelle : avantages et inconvénients de l'arborescence Règles d'utilisation d'un arbre pondéré". In : *Repère IREM no 15* (1994), p. 97–117.