

DOC
RES
A

INSTITUT DE RECHERCHE SUR L'ENSEIGNEMENT DES MATHÉMATIQUES

8716

IBC99001.PDF :

UNIVERSITÉ DE FRANCHE-COMTE



ARBRES ET PROBABILITES

IREM de LYON
BIBLIOTHEQUE

Université Claude Bernard -LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE Cedex

Jean-Pierre GRANGE



IREM de LYON
BIBLIOTHEQUE
Université Claude Bernard -LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE Cedex

ARBRES ET PROBABILITÉS

Jean-Pierre GRANGÉ
IREM de BESANCON

Ce travail a été réalisé grâce au soutien de la Direction de l'Enseignement Scolaire (DESCO) et de l'Institut Universitaire de Formation des Maîtres (I.U.F.M.) de Franche-Comté, avec le support logistique de la Direction de l'Enseignement Supérieur (D.G.E.S.), qu'ils en soient remerciés.

Je remercie également pour leurs relectures approfondies les membres du Groupe Probabilités de l'I.R.E.M. de Besançon : Marie-Jeanne BARTHÉLÉMY, Odile BAUMANN et Michel HENRY dont les conseils furent très précieux.

L'enseignement des probabilités au lycée, d'un volume relativement réduit, a cependant un objectif ambitieux : il est important de comprendre que, comme toute partie des mathématiques, **les probabilités s'inscrivent dans des modèles**. C'est à ce sujet que l'enseignant a la meilleure occasion de faire comprendre qu'un objectif fondamental des mathématiques est de proposer des modèles aux autres sciences. M. HENRY dans [1] p.78 définit un modèle comme *“une interprétation abstraite, simplifiée et idéalisée d'un objet du monde réel, ou d'un système de relations, ou d'un processus évolutif issu d'une description de la réalité”*. Celui-ci s'introduit souvent par l'intermédiaire d'une première abstraction de la réalité, parfois appelée « modèle pseudo-concret », porteuse d'hypothèses et de conventions indispensables à la construction du modèle probabiliste. Pour représenter ainsi certaines situations aléatoires concrètes, en particulier celles qui se ramènent à des situations d'équiprobabilité, on utilise une urne de Bernoulli¹ comme outil conceptuel de base. Les boules symbolisent les éléments qui peuvent être choisis au hasard et, d'après la situation réelle, ces choix s'effectuent selon un protocole déterminé. Pour mieux comprendre le déroulement du protocole expérimental d'une telle situation, il est très utile de s'aider d'un schéma, par exemple un arbre pondéré, appelé aussi « **arbre (du) probabiliste** ».

¹ Urne dont les boules sont de deux couleurs différentes pour symboliser un caractère dans une population.

Table des Matières

APPROCHE CARDINALISTE.....	7
APPROCHE FRÉQUENTISTE	11
<i>Expérience aléatoire à deux épreuves successives liées.....</i>	<i>12</i>
<i>Expérience aléatoire à deux épreuves répétées</i>	<i>13</i>
CONCEPTION FORMELLE.....	15
LES JEUNES ET LE CHÔMAGE.....	17
RÈGLES POUR LES ARBRES	21
<i>Premières règles de construction et d'utilisation dans les cas élémentaires</i>	<i>22</i>
<i>Règles complémentaires</i>	<i>23</i>
UN ÉNONCÉ, TROIS APPROCHES	25
<i>Première façon :</i>	<i>26</i>
<i>A l'aide d'un tableau :</i>	<i>27</i>
<i>A l'aide des arbres :</i>	<i>28</i>
<i>Explications :</i>	<i>29</i>
EXERCICES D'APPLICATIONS	31
<i>Exercice 1 : Les téléphones portables.....</i>	<i>31</i>
<i>Exercice 2 : Le temps en Probastatie.....</i>	<i>33</i>
<i>Exercice 3 : A la station-service</i>	<i>37</i>
CONCLUSION.....	41
BIBLIOGRAPHIE	43

APPROCHE CARDINALISTE

Si le professeur a plutôt une conception cardinaliste² de la probabilité, il choisira, pour introduire la notion de probabilité conditionnelle, une expérience aléatoire dont les événements sont construits par deux partitions d'une population, ce qu'on présente souvent sous forme d'un tableau de contingence comme ci-dessous :

Une enquête effectuée auprès des employés d'une entreprise concernant le sexe et le salaire a donné les résultats suivants :

	<i>salaire < 8 000 F</i>	<i>salaire ≥ 8 000 F</i>	
<i>femmes</i>	400	300	700
<i>hommes</i>	600	200	800
	1 000	500	1 500

Sur cette population, on considère :

une première expérience aléatoire E_1 : « Parmi les personnes employées on en choisit une au hasard » et on note P la probabilité relative aux événements de E_1 .

les événements A : « la personne gagne moins de 8 000 F » et B : « la personne est de sexe féminin ».

1°) Quelle est la probabilité que cette personne soit une femme ?

$$\text{C'est } P(B) = \frac{700}{1500} = \frac{7}{15}.^{(3)}$$

2°) Quelle est la probabilité que cette personne gagne moins de 8 000 F ?

$$\text{C'est } P(A) = \frac{1000}{1500} = \frac{10}{15}.$$

3°) Quelle est la probabilité que cette personne soit une femme gagnant moins de 8 000 F ?

L'événement « elle est une femme gagnant moins de 8 000 F » est représenté par $A \cap B$.

$$\text{Sa probabilité est } P(A \cap B) = \frac{400}{1500} = \frac{4}{15}.$$

² Voir A. TOTOHASINA dans [8]. Pour simplifier, c'est le fait de ramener systématiquement le calcul d'une probabilité au quotient de deux cardinaux.

³ L'hypothèse d'équiprobabilité, induite par l'expression « choix au hasard » dans l'énoncé, conduit à assimiler $P(B)$ à la proportion des femmes dans cette entreprise.

La suite du questionnement relatif à cet énoncé peut se faire de deux manières différentes qui correspondent à deux approches de la probabilité conditionnelle.

E₂ - Parmi les femmes, on en choisit une au hasard. Quelle est la probabilité qu'elle gagne moins de 8 000 F ?

On est dans la situation où il y a un **changement d'expérience aléatoire** (E₂), l'énoncé suggère que l'on a mis dans une urne fictive [B] les boules des employées de sexe féminin. La fonction probabilité associée à cette nouvelle expérience aléatoire est naturellement notée P_B, on peut l'appeler « probabilité sachant B ».

Le nouvel espace probabilisé est alors (B, P_B).

E'₂ - Sachant que c'est une femme, quelle est la probabilité qu'elle gagne moins de 8 000 F ?

On est dans la situation où il n'y a pas **changement de population** (même Ω), mais un renseignement sur les deux est connu et on s'intéresse à l'autre supposé ignoré ou perdu ou effacé ! Dans cette expérience E'₂, l'événement B a pour probabilité 1, on n'a donc plus la même fonction probabilité que dans l'expérience E₁. Ce renseignement connu induit un autre modèle de distribution de la probabilité sur Ω, on le note P_B. Le nouvel espace probabilisé est alors (Ω, P_B).

On a donc $P_B(A) = \frac{400}{700} = \frac{4}{7}$ et on constate que $P_B(A) = \frac{P(A \cap B)}{P(B)}$ ⁽⁴⁾, relation qui est

choisie comme définition de la probabilité conditionnelle :

Dans un espace probabilisé (Ω, P), étant donnés deux événements A et B tels que P(B) ≠ 0, on appelle "probabilité de A, lorsque B est réalisé" ou "probabilité sachant B, de A" le quotient :

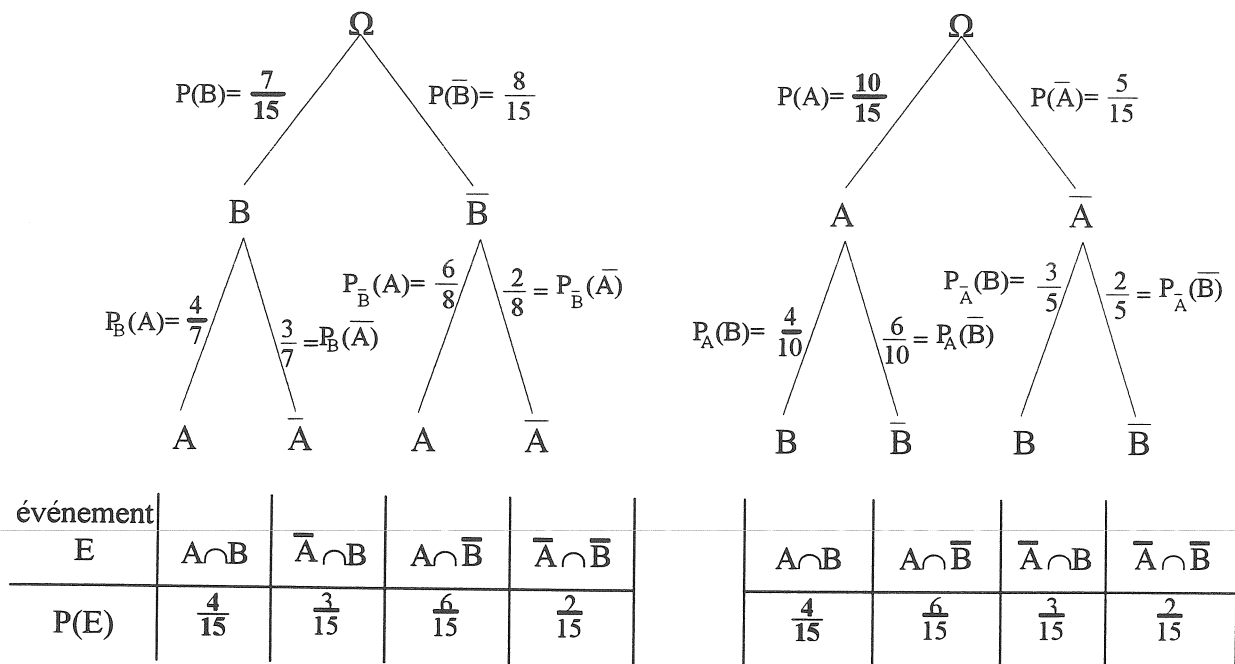
$$\frac{P(A \cap B)}{P(B)} \text{ noté } P_B(A) \text{ ou encore } P(A/B).$$

On en déduit immédiatement que : $P(A \cap B) = P(B) \times P_B(A)$.

Et de $P_A(B) = \frac{P(A \cap B)}{P(A)}$ on déduit : $P(A \cap B) = P(A) \times P_A(B)$.

Ces formules de calculs de la probabilité conjointe P(A ∩ B) par un produit de probabilités peuvent s'illustrer sur l'un ou l'autre des deux arbres probabilistes suivants :

⁽⁴⁾ Cette relation ne provient pas d'une situation numérique particulière ; elle est encore vraie lorsqu'il n'y a pas équiprobabilité. voir la démonstration de Michel HENRY dans [1] p 100, elle est également reprise p.15.



Cette approche cardinaliste est une conception réductrice de la probabilité qui attire beaucoup d'élèves mais ceux qui n'arriveront pas à surmonter cette approche éprouveront beaucoup de difficultés par la suite lorsque les probabilités ne se calculeront plus nécessairement par un quotient de deux cardinaux comme par exemple dans la loi de Poisson ou la loi Normale.

APPROCHE FRÉQUENTISTE

Rappelons d'abord la présentation par H. VENTSEL⁵ du sens fréquentiste donné à la probabilité : *“L'estimation numérique de l'éventualité d'un événement par la probabilité n'a de sens pratique que parce que les événements plus probables se produisent en moyenne plus souvent que les moins probables. Et si la pratique montre que, au fur et à mesure de l'augmentation du nombre d'expériences, la fréquence de l'événement a tendance à se stabiliser, s'approchant rigoureusement d'un nombre constant, il est naturel d'admettre que ce nombre est la probabilité de cet événement.”*

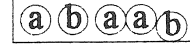
En classe de première, le professeur qui a une approche plutôt fréquentiste de la probabilité, a probablement étudié des expériences aléatoires à deux épreuves successives. Dans ces expériences le dessin d'un arbre permet la construction des événements élémentaires selon le protocole décrit dans l'énoncé. Au lycée, on distingue deux types différents de protocole qui servent de modèles, les tirages avec remise et les tirages sans remise ⁶ comme dans les situations suivantes.

⁵ Hélène VENTSEL, *Théorie des Probabilités*, ed. Mir, Moscou 1973.

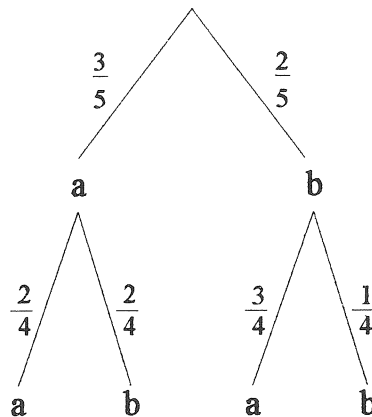
⁶ Il existe d'autres modèles comme celui de Polya où les tirages se font avec remise et rajout de boules identiques à la boule tirée. Ce protocole est réaliste pour étudier la contagion de maladies ou la contagion d'opinions.

Expérience aléatoire à deux épreuves successives liées

"On tire au hasard" ⁽⁷⁾ une lettre de l'urne ci-contre, on ne la remet pas puis on fait un second tirage.



Le déroulement des événements de cette expérience peut se représenter par l'arbre pondéré suivant :



cet arbre

nous conduit à

$$\Omega = \{ (a,a) ; (a,b) ; (b,a) ; (b,b) \}$$

Pour justifier l'attribution de probabilités aux événements élémentaires de cette expérience on peut expliquer le raisonnement suivant :

D'après l'approche fréquentiste de la probabilité, si on répète N fois cette expérience à deux épreuves liées, on peut **espérer** ⁽⁸⁾ que $\frac{3}{5}$ des N expériences donneront l'issue a au premier tirage puis $\frac{2}{4}$ de ces $\frac{3}{5} \times N$ expériences donneront b au second tirage donc $\frac{2}{4} \times \frac{3}{5}$ des N expériences donneront (a,b) . Ce qui permet d'**estimer** ⁽⁹⁾ $P\{(a,b)\}$ par $\frac{2}{4} \times \frac{3}{5} = \frac{3}{10}$.

Ce raisonnement conduit à cet énoncé : « Sur l'arbre, la probabilité du couple (a,b) s'obtient en multipliant les probabilités successives de a et de b » ⁽¹⁰⁾.

D'où la loi de probabilité associée à cette expérience à deux épreuves liées :

ω	(a,a)	(a,b)	(b,a)	(b,b)
$P(\{\omega\})$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{1}{10}$

Ce raisonnement précédent se généralise à une expérience comportant n épreuves aléatoires successives et conduit à cette sorte de "règle d'arbres" :

« Sur un arbre, la probabilité d'un n -uplet s'obtient en multipliant les probabilités rencontrées successivement sur les branches qui aboutissent à sa formation ».

⁽⁷⁾ Expression qui signifie, par contrat, qu'il y a équiprobabilité pour les éléments considérés.

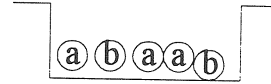
⁽⁸⁾ $\frac{3}{5} \times N$ est l'**espérance mathématique** de la variable aléatoire du nombre de réalisations de $\{a\}$ lorsqu'on répète N fois la première épreuve.

⁽⁹⁾ C'est le meilleur **estimateur** d'après la théorie statistique de l'Estimation.

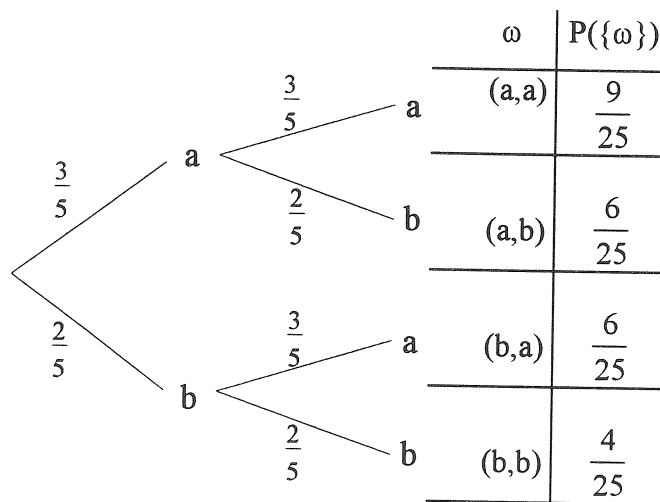
⁽¹⁰⁾ En réalité cet énoncé représente la propriété selon laquelle la probabilité conjointe $P(A_1 \cap B_2)$ est égale au produit des probabilités $P(A_1) \times P_{A_1}(B_2)$, où A_i et B_i sont les événements définis page suivante.

Expérience aléatoire à deux épreuves répétées

On tire au hasard une lettre de l'urne ci-contre, on la note, on la remet dans l'urne puis on fait un second tirage.

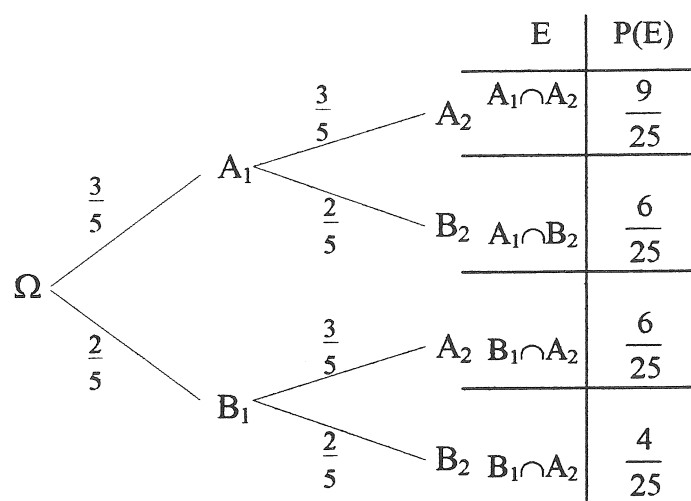


Le déroulement des événements de cette expérience peut se représenter par l'arbre pondéré ci-dessous et, comme on peut faire le même raisonnement que précédemment en répétant N de ces expériences, la même "règle d'arbre"⁽¹¹⁾ précédente s'applique et on obtient la loi de probabilité associée à cette expérience à deux épreuves répétées :



Les arbres pondérés utilisés dans ces deux exemples sont des arbres de construction des événements élémentaires. **Ils ne sont pas encore des arbres probabilistes** mais en forment une bonne approche. Par la suite, l'arbre probabiliste qui représentera cette dernière expérience est le suivant où figure l'univers Ω ainsi que les événements A_i : « on obtient a au $i^{\text{ème}}$ tirage » et B_i : « on obtient b au $i^{\text{ème}}$ tirage »

⁽¹¹⁾ En réalité ici, cette "règle d'arbres" représente la propriété d'indépendance stochastique qui se traduit par un produit de probabilités.



On fait alors constater aux élèves que dans ces deux expériences aléatoires la probabilité conjointe $P(A_1 \cap B_2)$ est égale au produit de deux probabilités dont l'une est $P(A_1)$ et l'autre une nouvelle probabilité, celle de B_2 sachant que A_1 est réalisé, que l'on note : $P_{A_1}(B_2)$. Ainsi dans un cas général de deux événements A et B d'un univers Ω on admet que : $P(A \cap B) = P(A) \times P_A(B)$ ou encore, si $P(A) \neq 0$: $P_A(B) = \frac{P(A \cap B)}{P(A)}$

Contrairement à la première introduction, ces deux expériences sont liées au temps par des épreuves successives et peuvent renforcer les conceptions chronologiste ou causaliste¹² de cette probabilité. Si elles sont trop prégnantes, ces conceptions risquent de présenter des obstacles à la réversibilité de la probabilité conditionnelle lorsqu'il s'agira de déterminer les probabilités de causes ayant observé des conséquences (cf M. HENRY dans [6] p.97-98 ou J.P. GRANGÉ dans [1] p.366). Il faudrait alors faire suivre ces deux expériences par une situation ensembliste présentant deux caractères non hiérarchisés comme le sexe et le salaire des employés d'une entreprise¹³. Cette situation aléatoire permet la construction de deux arbres.

¹² Conception causaliste très simplifiée :

l'événement conditionnant est toujours considéré comme une cause de l'événement conditionné.

Conception chronologiste très simplifiée :

l'événement conditionnant est toujours considéré comme antérieur à l'événement conditionné.

On trouvera davantage de précisions sur ces conceptions chez A. TOTOHASINA dans [8].

¹³ Voir p. 7.

CONCEPTION FORMELLE

Le professeur qui se place d'emblée dans le modèle probabiliste et donne directement la définition de la probabilité conditionnelle fait une présentation formelle. Cette introduction ne comporte sans doute pas le risque d'accentuer des conceptions réductrices, mais son caractère abstrait est peu acceptable pour nos élèves. En réalité, dans la théorie probabiliste, cette définition découle de la construction du modèle dont M. HENRY donne la présentation suivante dans [1] p.100 en respectant l'esprit des programmes actuels de lycée qui proposent une compréhension fréquentiste de la notion de probabilité :

« Notons : Ω l'ensemble référentiel, représentant l'expérience aléatoire primitive $\mathcal{E}$ et P la distribution de probabilité sur Ω .

B une partie de Ω de probabilité non nulle, représentant un "événement" au sens familier associé à $\mathcal{E}$ et noté $\mathcal{B}$. En théorie des probabilités, B est encore appelé "événement de Ω "

A un événement de Ω représentant un événement au sens familier associé à $\mathcal{E}$ noté $\mathcal{A}$ et dont on cherche la probabilité.

Si l'on reproduit $\mathcal{E}$ un grand nombre n de fois, la valeur observée $f_{\mathcal{B}}$ de la fréquence des réalisations de $\mathcal{B}$ sera proche autant qu'on le veut de $P(B)$:

$$f_{\mathcal{B}} = \frac{\text{nombre de réalisations de } \mathcal{B}}{n}$$

De même $P(A \cap B)$ sera approché par : $f_{\mathcal{A} \cap \mathcal{B}} = \frac{\text{nombre de réalisations conjointes de } \mathcal{A} \text{ et } \mathcal{B}}{n}$

Si l'on s'intéresse à la fréquence des réalisations de $\mathcal{A}$ parmi les issues qui réalisent aussi $\mathcal{B}$, celle-ci est donnée par :

$$f_{\mathcal{A}|\mathcal{B}} = \frac{\text{nombre de réalisations conjointes de } \mathcal{A} \text{ et } \mathcal{B}}{\text{nombre de réalisations de } \mathcal{B}}, \text{ d'où : } f_{\mathcal{A}|\mathcal{B}} = \frac{f_{\mathcal{A} \cap \mathcal{B}}}{f_{\mathcal{B}}}.$$

La probabilité d'obtenir A , lorsque l'on considère l'expérience aléatoire $\mathcal{E}_{\mathcal{B}}$ consistant à n'accepter un résultat de $\mathcal{E}$ que lorsque $\mathcal{B}$ est réalisé, est estimée par $f_{\mathcal{A}|\mathcal{B}}$. On la note $P_B(A)$.

La relation obtenue entre les fréquences conduit donc à la relation liant cette nouvelle probabilité de A , aux probabilités a priori de $A \cap B$ et de B : $P_B(A) = \frac{P(A \cap B)}{P(B)}$. »

Ainsi, dans une présentation formelle, le professeur écrit :

Définition :

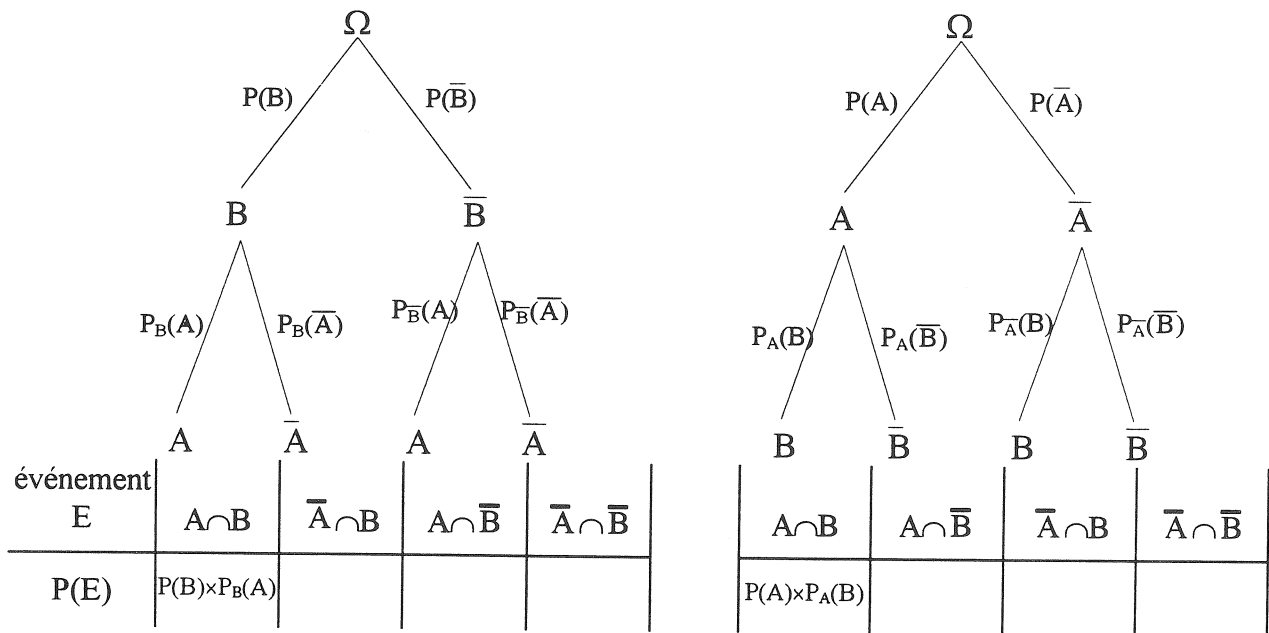
Dans un espace probabilisé (Ω, P) , étant donnés deux événements A et B tels que $P(B) \neq 0$, on appelle "probabilité de A , sachant que B est réalisé" ou "probabilité sachant B , de A " le quotient :

$$\frac{P(A \cap B)}{P(B)} \text{ noté } P_B(A) \text{ ou encore } P(A/B).$$

Cette définition signifie aussi que si $P(A) \neq 0$, $P_A(B) = \frac{P(A \cap B)}{P(A)}$ ¹⁴

De ces deux écritures, on tire : $P(A \cap B) = P(B) \times P_B(A)$.
et $P(A \cap B) = P(A) \times P_A(B)$.

Formules que l'on peut représenter sur l'un ou l'autre des deux arbres probabilistes suivants :



Il reste alors à donner aux élèves des exercices d'application comportant des cardinaux, des épreuves successives mais surtout des pourcentages car ce sont eux qui mettent en acte la notion de probabilité.

Dans l'exercice qui suit, où interviennent des pourcentages, on va constater combien les arbres probabilistes sont des outils de calculs mais aussi des outils de raisonnement et de démonstration.

¹⁴ On peut ainsi considérer trois distributions de probabilités sur Ω : P , P_B et P_A qui sont liées par ces relations. Dans ce cas, ces écritures s'autorisent l'abus de considérer que Ω représente successivement les trois expériences aléatoires $\mathcal{E}$, $\mathcal{E}_B$ et $\mathcal{E}_A$.

LES JEUNES ET LE CHÔMAGE

Dans un pays où on sait que 30% des jeunes sont des chômeurs et que 40% des chômeurs sont des jeunes, un individu jeune a-t-il plus de chances d'être chômeur qu'un autre individu ?

Pour répondre à une telle question posée dans le langage familier, il est indispensable d'avoir une définition précise des termes utilisés d'une part et de préciser le modèle probabiliste choisi d'autre part. Pour cette question, supposons les définitions suivantes :

"un individu jeune" est un individu âgé de moins de 30 ans à la date du sondage ;

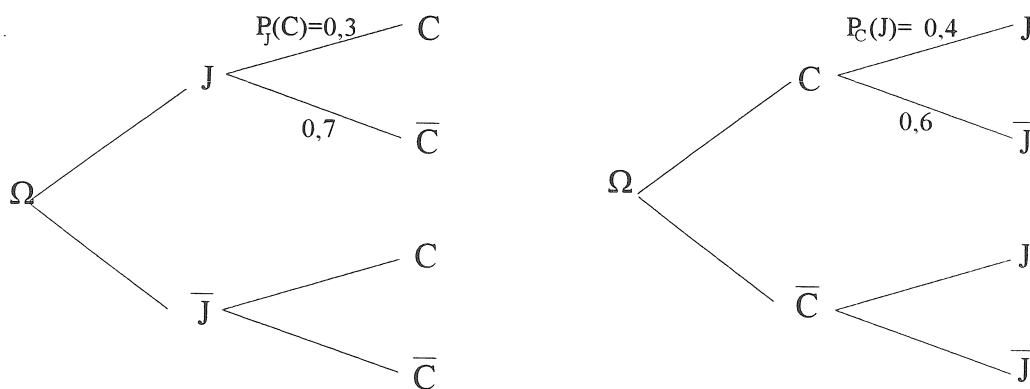
"un individu chômeur" est un individu demandeur d'emploi à la date du sondage ;

"la population active" est l'ensemble des travailleurs et des demandeurs d'emploi.

On décide ensuite que :

l'expérience aléatoire consiste à choisir un individu au hasard¹⁵ dans la population active de ce pays. La probabilité que cet individu soit jeune peut donc être estimée par la proportion des jeunes dans la population. De même, si on sait que c'est un chômeur la probabilité qu'il soit jeune est estimée par la proportion des jeunes parmi les chômeurs, c'est à dire 40 %. De ces données on déduit le modèle probabiliste suivant : soit Ω l'univers représentant le choix aléatoire d'un individu dans la population, les conditions de l'expérience conduisent au modèle d'équiprobabilité. Dans ce modèle mathématique, notons P cette probabilité mise en place sur Ω , notons J l'événement « il est jeune » et C l'événement « il est chômeur ». L'énoncé conduit à prendre : $P_J(C) = 0,3$ et $P_C(J) = 0,4$

Pour aider la compréhension de l'énoncé et le raisonnement à mettre en oeuvre, on dessine les deux arbres probabilistes qui sont particulièrement bien adaptés lorsqu'une expérience aléatoire est décrite par des probabilités conditionnelles :



Lorsqu'on connaît quelques règles d'arbres¹⁶, en particulier celles concernant le passage d'un arbre à l'autre, on s'aperçoit alors que les données sont insuffisantes pour calculer quelque autre probabilité donc pour répondre à la question posée. Voilà un exemple où les arbres probabilistes servent à faire un raisonnement.

¹⁵ Ce qui signifie que tous les éléments de la population ont la même "chance" d'être choisis.

¹⁶ Voir p.21.

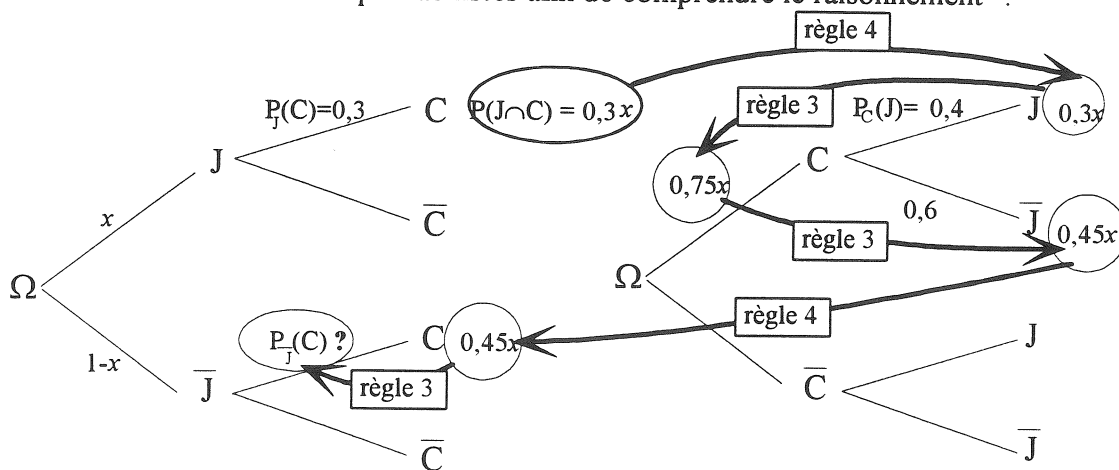
Notons alors x la proportion des individus jeunes dans ce pays et, par exemple, cherchons x pour que les événements J et C soient indépendants, c'est à dire, pour qu'un individu jeune n'ait ni plus ni moins de chances d'être chômeur qu'un autre individu.

Compte tenu du choix de l'équiprobabilité pour le modèle, on a $P(J) = x$ et, $P(C)$ étant supposé différent de 0, les événements J et C sont indépendants si et seulement si : $P(J) = P_C(J)$ ⁽¹⁷⁾, c'est à dire $x = 0,4$.

Ce résultat correspond bien à l'idée simple qu'on peut se faire de l'indépendance dans le cas d'équiprobabilité : la proportion des jeunes par rapport à la population active est égale à celle des jeunes parmi les chômeurs. Intéressons-nous alors à la question suivante :

Si cette proportion de jeunes (x) est supérieure à 40%, un individu jeune a-t-il plus de chances qu'un autre d'être chômeur ? Rappelons que, dans ce pays, 30% des jeunes sont des chômeurs et que 40% des chômeurs sont des jeunes.

Il convient donc de comparer $P_I(C)$ et $P_J(C)$ et pour répondre à cette question, on peut s'aider des flèches liant les deux arbres probabilistes afin de comprendre le raisonnement¹⁸.



On peut bien sûr se passer des arbres, mais il est nécessaire de justifier :

$$P_{\bar{J}}(C) = \frac{P(\bar{J} \cap C)}{P(\bar{J})} \text{ avec } P(\bar{J}) = 1 - x, \text{ puis deux raisonnements peuvent convenir.}$$

$$P(\bar{J} \cap C) = P(C) \times P_c(\bar{J}) = P(C) \times [1 - P_c(J)].$$

$$\text{Or, } P_C(J) = \frac{P(J \cap C)}{P(C)} = \frac{P(J) \times P_J(C)}{P(C)},$$

d'où on tire $P(C) = \frac{P(J) \times P_J(C)}{P_C(J)}$ ($= 0,75x$) et

$$P(\bar{J} \cap C) = \frac{P(J) \times P_J(C)}{P_C(J)} \times [1 - P_C(J)] (= 0,45x) \quad 19$$

$$P(\bar{J} \cap C) = P(C) - P(J \cap C) = \frac{P(J \cap C)}{P_C(J)} - P(J \cap C)$$

$$\text{d'où } P(\bar{J} \cap C) = P(J \cap C) \times \left[\frac{1}{P_C(J)} - 1 \right]$$

$$\text{d'où } P(\bar{J} \cap C) = P(J) \times P_J(C) \times \left[\frac{1 - P_C(J)}{P_C(J)} \right]^{20}$$

et finalement :
$$P_{\bar{J}}(C) = \frac{P(J) \times P_J(C)}{P_C(J)} \times \frac{[1 - P_C(J)]}{P(\bar{J})} = \frac{x \times 0,3 \times 0,6}{0,4 \times (1-x)} = \frac{0,45x}{1-x}$$

¹⁷ Relation parfois donnée comme définition de deux événements indépendants, voir par exemple J.P. GRANGÉ dans [1] p. 361.

¹⁸ Des règles d'utilisation des arbres probabilistes sont précisées p.21.

¹⁹ Ce raisonnement correspond au parcours fléché ci-dessus.

²⁰ Ce raisonnement correspond plutôt à une approche ensembliste.

$P_{\bar{J}}(C)$ est une fonction homographique de x , strictement croissante sur $[0 ; 1[$, donc :

si $x > 0,4$ alors $P_{\bar{J}}(C) > \frac{0,45 \times 0,4}{1-0,4} = 0,3 = P_J(C)$.

Ce sont alors les non jeunes qui ont la plus grande probabilité d'être chômeurs dès que $x > 0,4$.

On peut penser que tous ces calculs sont inutiles pour obtenir cette conclusion, car on sait que 40% des chômeurs sont des jeunes et que les jeunes représentent plus de 40% de la population active, ce qui montre qu'il y a effectivement une proportion plus importante des non jeunes parmi les chômeurs que parmi la population active, et d'autre part le renseignement 30% des jeunes sont des chômeurs est alors inutile. Mais le calcul précédent a l'avantage de donner la valeur de cette probabilité pour toute valeur connue x de la proportion des jeunes dans cette population. Ainsi pour $x = 0,5$ on obtient : $P_{\bar{J}}(C) = 0,45$ alors que $P_J(C)$ est toujours égale à 0,3. Et si $x = 0,45$ par exemple, alors $P_{\bar{J}}(C) \approx 0,37$.

Voilà un exemple où **les arbres probabilistes aident à faire des calculs** grâce aux nombres écrits sur les branches et les feuilles. **Ces arbres probabilistes servent aussi à élaborer une démonstration** grâce aux formules écrites ou sous entendues sur les branches et les feuilles. Il paraît donc important lors d'un apprentissage que les nombres écrits sur les branches soient toujours accompagnés de leurs significations mathématiques, comme par exemple « $P_J(C) = 0,3$ ».

Les deux arbres précédents qui ont servi de support au raisonnement et au calcul, suivent des règles de construction et d'utilisation qui découlent de la théorie probabiliste. Dans des cas simples, on peut indiquer ces règles aux élèves mais il y a un risque que les didacticiens appellent le glissement métacognitif, c'est à dire qu'on se limite à l'enseignement de ces règles et au mode d'emploi des arbres à la place des définitions et propriétés mathématiques²¹. La présence des règles suivantes a donc, ici, pour objet de proposer aux enseignants en probabilités, une explication des principes de construction de l'arbre probabiliste, quant à l'utilisation des arbres, on va voir qu'on se heurte à certaines limites. L'enjeu pour les élèves est de pouvoir être aidés par ce schéma afin de mieux comprendre des notions probabilistes de base.

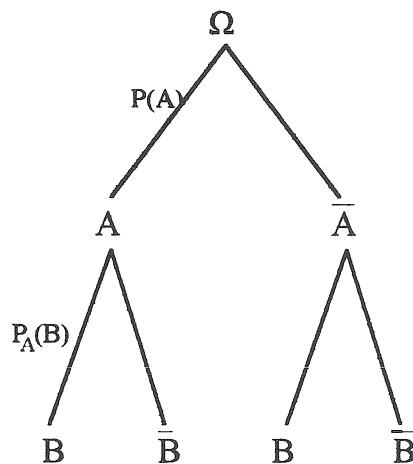
²¹ De la même façon qu'on limiterait l'enseignement de la proportionnalité aux tableaux de proportionnalité.

RÈGLES POUR LES ARBRES

Ceci est une contribution à la mise en forme des principales règles de construction et d'utilisation des arbres probabilistes ²³.

La modélisation d'une expérience aléatoire dont les événements considérés sont en nombre fini et pas très nombreux peut se représenter par un arbre mais celui-ci **induit une conception chronologiste** ou causaliste même lorsqu'elle n'existe pas dans l'expérience aléatoire. Ainsi lorsqu'on utilise les arbres comme outil de résolution, il faut être très conscient de ces conceptions réductrices qu'ils portent en eux-mêmes.

Il est important que l'utilisation de l'outil arbre soit progressive et donc que les premiers exemples rencontrés soient du type élémentaire suivant. C'est lui qui permet de mettre en place et de retrouver les théorèmes de probabilités tels qu'ils sont énoncés dans le cours.

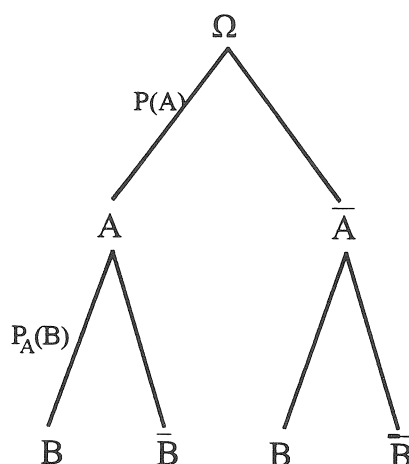


On peut distinguer deux séries de règles, une pour s'appropriier l'arbre-outil dans les cas élémentaires et une autre pour le perfectionner dans les cas plus complexes.

²³ B. PARZYSZ distingue des règles de traitement et de conversion des arbres dans [1] p.233

Premières règles de construction et d'utilisation dans les cas élémentaires

1 - Un **arbre** se fabrique et se lit de la gauche vers la droite ou du haut vers le bas. Il est formé de branches et de feuilles. L'origine de l'arbre est l'événement certain, l'univers Ω sur lequel est définie une probabilité qui sera notée P .



2 - Une **branche** représente un lien probabiliste entre un événement A , considéré comme antécédent, et un événement B considéré comme succédant à A . La probabilité notée sur cette branche est la probabilité de B conditionnée par la réalisation de l'événement A .

$$A \xrightarrow{P_A(B)} B, \text{ cas particulier : } P_{\Omega}(A) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A).$$

3 - Une succession de plusieurs branches avec leurs événements est un **chemin**. Au bout d'un chemin se trouve une **feuille**, pas toujours indiquée, qui représente l'événement conjoint des événements rencontrés sur ce chemin et sa probabilité est égale au produit des probabilités notées sur chacune de ses branches.

$$\Omega \xrightarrow{P(A)} A \xrightarrow{P_A(B)} B \quad \begin{array}{|c|c|} \hline \text{événement} & \text{probabilité} \\ \hline A \cap B & P(A \cap B) = P(A) \times P_A(B) \\ \hline \end{array}$$

4 - La conjonction, notée $A \cap B$, de deux événements A et B étant commutative, un chemin n'est pas ordonné.

Par exemple, on peut écrire : $P(A \cap B) = P(A) \times P_A(B) = P(B \cap A) = P(B) \times P_B(A)$.

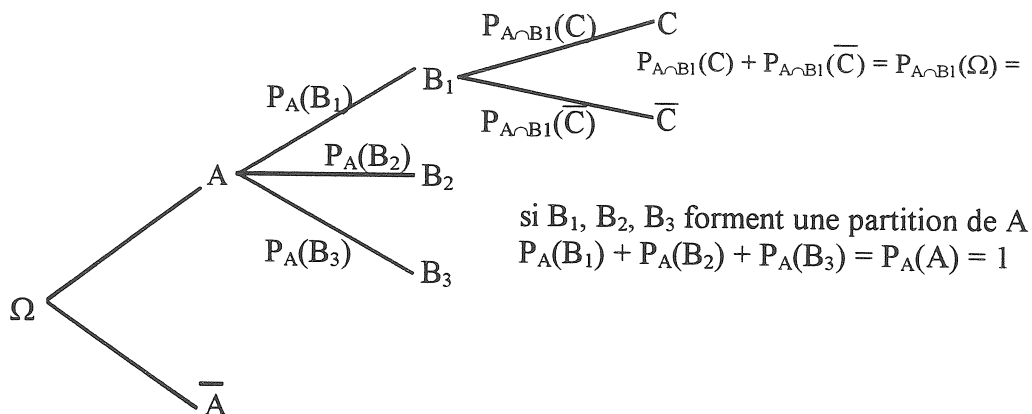
Ainsi lorsque deux ou plusieurs arbres traduisent la même situation aléatoire, la probabilité relative à une feuille peut être transférée d'un arbre à l'autre ou aux autres²³.

Pour illustrer des expériences aléatoires moins élémentaires on peut ajouter des règles complémentaires comme les suivantes :

²³ La règle 4 est souvent utilisée dans des situations de passage d'un arbre à l'autre encore appelées situations de retournement (ou d'inversion) de l'arbre probabiliste, voir p.18. Elle permet d'illustrer la formule de Bayes.

Règles complémentaires

5 - Sur un arbre, si les événements succédants à un événement A en forment une partition, alors la somme des probabilités notées sur les branches qui suivent A est égale à 1.



De cet arbre, on peut extraire le chemin qui mène à l'événement-feuille $A \cap B_1 \cap C$:

					événement E	probabilité P(E)
Ω	$\xrightarrow{P(A)}$	A	$\xrightarrow{P_A(B_1)}$	B_1	$\xrightarrow{P_{A \cap B_1}(C)}$	C
					$A \cap B_1 \cap C$	$P(A) \times P_A(B_1) \times P_{A \cap B_1}(C).$

6 - On dit qu'un arbre est **complet** si, pour chacun de ses événements, les événements succédants en forment une partition.

Sur un arbre complet, l'ensemble des chemins qui joignent Ω à un événement final représente des événements qui forment une partition de Ω , ainsi la somme de leurs probabilités vaut 1.

Par exemple sur l'arbre ci-dessus, lorsque B_1, B_2, B_3 forment une partition de A, on a :

$$P(A \cap B_1 \cap C) + P(A \cap B_1 \cap \bar{C}) + P(A \cap B_2) + P(A \cap B_3) + P(\bar{A}) = 1$$

Ces dernières règles donnent un aperçu des difficultés qui vont surgir si on veut progresser dans la voie des arbres probabilistes alors que dans la voie mathématique les théorèmes probabilistes comme les définitions sont rigoureux et décontextualisés, c'est donc vers elle qu'il faut peu à peu se diriger. En ce qui concerne les arbres, il serait donc préférable de s'en tenir aux cas simples et ne pas croire que celui-ci est un remède à toutes les difficultés rencontrées dans l'enseignement des probabilités. Comme toute représentation, il a ses limites et il engendre ou renforce des conceptions réductrices²⁵, principalement la conception chronologiste. Il existe d'autres représentations, comme les tableaux à double entrée qui, eux aussi sont une aide efficace à la compréhension de certaines notions en probabilités, mais qui, eux aussi ont leurs limites et leurs conceptions réductrices.

Sur la base de l'exercice de probabilité du Baccalauréat série S, de Nouvelle Calédonie du mois de novembre 1997, regardons comment pouvaient intervenir ces différentes approches et leurs représentations éventuelles.

²⁵ Voir un exercice vécu par J.-P. GRANGÉ dans [4].

UN ÉNONCÉ, TROIS APPROCHES

D'après l'exercice commun à tous les candidats du Baccalauréat S., sujet de Nouvelle Calédonie, novembre 1997.

Un artisan est contacté à domicile par ses clients sur appel téléphonique et dispose d'un répondeur. On a constaté que chaque jour entre 11 heures et midi :

- a) Quand l'artisan est absent, il branche systématiquement le répondeur.*
- b) Quand il est présent, il le branche une fois sur trois.*
- c) Quand un client téléphone, il a quatre chances sur cinq d'obtenir le répondeur et une chance sur cinq d'obtenir l'artisan.*

Un client téléphone à l'artisan un jour entre 11 heures et midi.

Si on veut aider les élèves à s'approprier cette situation en tant que situation aléatoire on peut poser la question suivante qui est tout à fait dans l'esprit du programme des classes de Premières :

Indiquer les issues possibles concernant cette expérience aléatoire.

Pour les correcteurs, il est ensuite préférable d'imposer des notations :

On note P la fonction probabilité associée à cette expérience²⁶ et on note :

R l'événement : « le client obtient le répondeur »

A l'événement : « l'artisan est présent ».

Pour tout événement X , on note $\bar{X}$ l'événement contraire de X .

L'objectif est de faire trouver aux élèves la probabilité que l'artisan soit présent et on peut s'y prendre de plusieurs façons :

²⁶ Cette précision est indispensable comme on écrit : « on note f la fonction définie par ». Sans cette précision, la lettre P a tendance à être perçue comme une abréviation du mot probabilité, d'autant plus que l'événement de Ω auquel elle s'applique a lui aussi tendance à être confondu avec l'événement au sens familier.

Première façon :

celle utilisée par l'auteur de l'énoncé initial.

- 1°) Déterminer la probabilité $P(R)$, ainsi que les probabilités conditionnelles $P_A(R)$ et $P_{\bar{A}}(R)$.
 2°) Exprimer $P(R)$ en fonction de $P_A(R)$, $P_{\bar{A}}(R)$ et $P(A)$.
 3°) En déduire l'égalité $\frac{4}{5} = -\frac{2}{3}P(A) + 1$ et calculer la probabilité que l'artisan soit présent.

Ce cheminement choisi pour obtenir $P(A)$ semble prévu pour être résolu sans l'aide de représentations mais principalement en exploitant, grâce à une certaine expérience, les formules de probabilités du programme. Les quelques résultats dont je disposais au moment des corrections montrent que ceci évalue principalement une bonne mémoire et une bonne aptitude au calcul algébrique mais pas nécessairement une bonne compréhension de la notion de probabilité conditionnelle.

Solution attendue :

1°) D'après l'énoncé, on a : $P(R) = \frac{4}{5}$; $P_A(R) = \frac{1}{3}$ et $P_{\bar{A}}(R) = 1$.

2°) D'après la formule des probabilités totales : $P(R) = P(R \cap A) + P(R \cap \bar{A})$ et d'après les probabilités conditionnelles, on obtient : $P(R) = P(A) \times P_A(R) + P(\bar{A}) \times P_{\bar{A}}(R)$ avec $P(\bar{A}) = 1 - P(A)$

3°) En utilisant les résultats du 1°), l'égalité précédente devient : $\frac{4}{5} = P(A) \times \frac{1}{3} + [1 - P(A)] \times 1$.

D'où on déduit $\frac{4}{5} = -\frac{2}{3}P(A) + 1$ soit $P(A) = -\frac{1}{5} \times (-\frac{3}{2}) = \frac{3}{10}$

A l'aide d'un tableau :

- 1°) Déterminer la probabilité $P(R)$, ainsi que les probabilités conditionnelles $P_A(R)$ et $P_{\bar{A}}(R)$.
 2°) On pose $P(A) = x$, exprimer en fonction de x les probabilités $P(A \cap R)$ et $P(\bar{A} \cap R)$.
 3°) En déduire $P(A)$

Ce cheminement choisi pour obtenir $P(A)$ semble prévu pour être résolu avec l'aide d'un tableau (diagramme de Carroll) car dès qu'on propose au candidat de poser la probabilité marginale $P(A) = x$, on lui permet de compléter deux cases intérieures du tableau et, en appliquant la formule des probabilités totales, de remplir ce tableau en entier donc de trouver $P(A)$. Dans cette situation, cette démarche est très proche de la précédente car elle nécessite aussi plusieurs questions intermédiaires.

Solution attendue :

1°) D'après l'énoncé, on a : $P(R) = \frac{4}{5}$; $P_A(R) = \frac{1}{3}$ et $P_{\bar{A}}(R) = 1$.

2°) Si on pose $P(A) = x$ on obtient le tableau suivant où l'on a complété en caractères gras la probabilité de l'événement contraire $\bar{A}$ ainsi que les probabilités conjointes :

$$P(A \cap R) = P(A) \times P_A(R) = x \times \frac{1}{3} \text{ et } P(\bar{A} \cap R) = P(\bar{A}) \times P_{\bar{A}}(R) = (1 - x) \times 1 = 1 - x$$

	A	$\bar{A}$	
R	$\frac{1}{3}x$	$1 - x$	$\frac{4}{5}$
$\bar{R}$			$\frac{1}{5}$
	x	$1 - x$	1

3°) D'après les résultats précédents, la formule des probabilités totales $P(R) = P(A \cap R) + P(\bar{A} \cap R)$ devient $\frac{4}{5} = \frac{1}{3}x + 1 - x$, ce qui donne $-\frac{1}{5} = -\frac{2}{3}x$ soit $x = \frac{3}{10}$

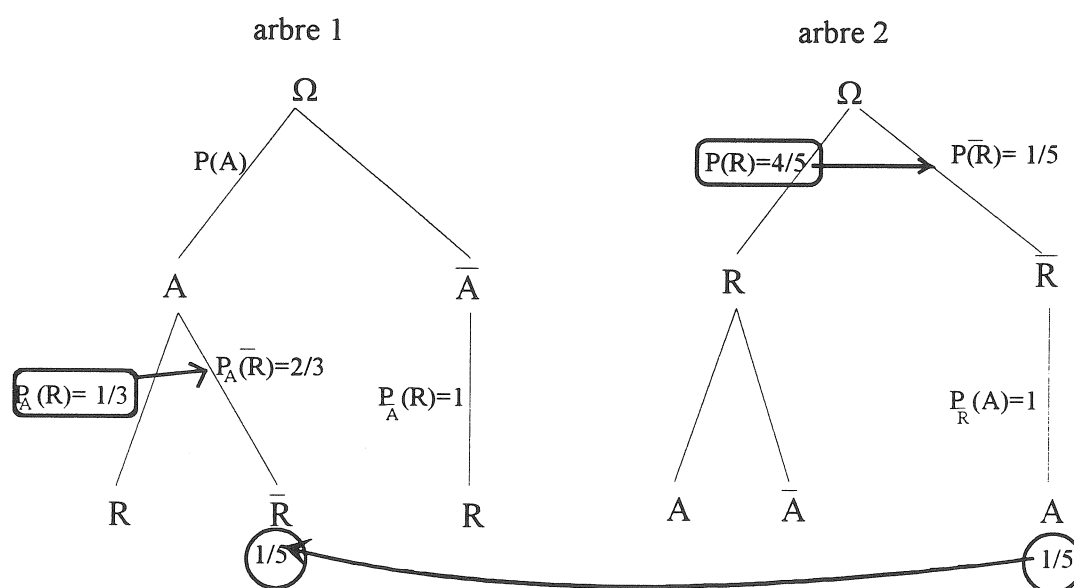
A l'aide des arbres :

1°) Dessiner les deux arbres possibles décrivant cette expérience aléatoire en y faisant figurer les probabilités que l'énoncé permet de connaître, par exemple $P(R) = \frac{4}{5}$.

2°) Calculer $P(\bar{R} \cap A)$, en déduire $P(A)$.

Ce cheminement qui impose les arbres pondérés comme outil de résolution, est très directif, mais il est tout à fait dans l'esprit de ce que pourrait être un début d'énoncé de Baccalauréat à partir de la session 1999. La représentation en arbres est tout à fait justifiée ici, car on reconnaît dans l'énoncé des renseignements se traduisant par des probabilités conditionnelles :

Solution attendue :



Le résultat $1/5$ obtenu sur le second arbre s'explique par : $P(\bar{R} \cap A) = P(\bar{R}) \times P_{\bar{R}}(A) = 1/5$

D'où la réponse, grâce au premier arbre, en expliquant :

$$\begin{aligned} P(\bar{R} \cap A) &= 1/5 \\ P(A) \times P_A(\bar{R}) &= 1/5 \\ P(A) \times 2/3 &= 1/5 \\ P(A) &= 3/10 \end{aligned}$$

Regardons un peu, comment les trois approches de la probabilité décrites p.7, p.11 et p.15 sont à l'origine de ces différents cheminements ?

Explications :

La formalisation est, comme dans tout domaine mathématique, la finalité de cet enseignement des probabilités conditionnelles mais elle n'est peut-être pas le moyen le mieux adapté pour son apprentissage. Elle correspond exactement à la définition telle qu'elle est imposée par les programmes : $P_B(A) = \frac{P(A \cap B)}{P(B)}$. C'est une définition d'essence ensembliste, elle rassure car elle fait penser à un quotient et rappelle, à tort, la formule de Laplace dans le cas d'équiprobabilité $\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$. Elle se heurte très vite au vide de connaissances en théorie des ensembles des élèves actuels auxquels on fait croire que $A = (A \cap B) \cup (A \cap \bar{B})$ est une évidence, que l'on justifie parfois comme on peut à partir de la situation concrète de l'énoncé. Il en est de même pour $(A \cup B) \cap B = B$ et que dire des lois de De Morgan telles que $\overline{A \cup B} = \bar{A} \cap \bar{B}$ pour lesquelles le retour au diagramme de Venn est le seul recours de l'enseignant mais convainc-il les élèves ?

Le tableau est une représentation assez fréquemment choisie par les enseignants parce que c'est un peu Venn ou plus exactement Karnaugh dans la théorie des ensembles et qu'on le rencontre dans différentes parties des mathématiques comme diagramme de Veitch en algèbre de Boole et comme tableau de contingence et diagramme de Carroll en statistiques et probabilité. Ce dernier diagramme est fondé sur la formule des probabilités totales ²⁷: $P(A \cap B) + P(A \cap \bar{B}) = P(A)$, formule additive, donc très simple d'approche pour les élèves. Elle correspond à la compréhension suivante de la probabilité conditionnelle :

$$P_B(A) = \frac{P(A \cap B)}{P(A \cap B) + P(A \cap \bar{B})} \quad \text{qui se traduit dans le cas où il y a équiprobabilité par}$$

$$P_B(A) = \frac{\text{card}(A \cap B)}{\text{card}(A \cap B) + \text{card}(A \cap \bar{B})}. \quad \text{Son avantage est de ne pas être séquentiel, c'est à dire de ne pas induire de chronologie entre certains événements et ses inconvénients majeurs sont :}$$

1°) Il renforce la conception cardinaliste par le calcul systématique d'une probabilité par une fraction.

2°) On ne peut y faire figurer d'une manière simple et pratique des probabilités conditionnelles, ce qui est particulièrement gênant lorsqu'on veut faire comprendre et utiliser cette notion.

²⁷ La formule des probabilités totales pour un système complet d'événements n'est plus au programme officiel des classes Terminales S à la rentrée 98, mais dans le cas particulier de deux événements et de leurs complémentaires, il semble impossible de s'en passer car c'est la justification des tableaux qui disparaîtrait.

L'arbre probabiliste est la représentation la plus naturelle pour exploiter des probabilités conditionnelles. La plus naturelle car, dans des situations aléatoires chronologistes ou causalistes, il retrace la succession des événements comme ils sont décrits dans l'énoncé. L'arbre correspond davantage à l'écriture suivante qui donne la probabilité conjointe : $P(A \cap B) = P(B) \times P_B(A)$ où la probabilité conditionnelle $P_B(A)$ apparaît comme coefficient multiplicateur inférieur à 1, donc réducteur, ce qui est logique puisqu'elle intervient pour calculer la probabilité de $A \cap B$ qui est une partie de B.

Il convient également aux situations ensemblistes grâce auxquelles on comprend très bien l'existence de plusieurs arbres pour une même situation aléatoire. Son inconvénient est que, dans des situations chronologistes ou causalistes, il empêche justement de concevoir le second arbre. Ceci correspond à la difficulté épistémologique que rencontre certains élèves lorsqu'on leur pose la question de déterminer la probabilité d'une cause ayant observé une conséquence qui a eu lieu après cette cause.

L'exemple précédent montre comment notre approche personnelle des probabilités influence le questionnement qu'on propose, oriente vers la représentation éventuellement attendue et donc impose la démarche vers la solution. Regardons dans les trois exercices suivants, quelle démarche semble la mieux appropriée pour les solutionner.

EXERCICES D'APPLICATIONS

Exercice 1

Extrait de l'exercice commun à tous les candidats du Baccalauréat E.S., sujet national, juin 1998.

Un magasin de distribution vend deux types de téléphones portables : des téléphones standard et des téléphones miniatures. Il propose aussi deux types d'abonnements mensuels : l'abonnement d'une heure et l'abonnement 2 h 30.

Le service marketing effectue une enquête sur un échantillon de 2 000 clients ayant acheté dans ce magasin, pendant l'année en cours, un téléphone et un seul de l'un des types vendus et ayant opté pour un seul des abonnements proposés.

Sur les 2 000 clients interrogés, 1 200 ont acheté le modèle standard.

Sur ces 2 000 clients, 960 ont choisi l'abonnement d'une heure.

Un client est pris au hasard dans l'échantillon. On note les événements :

S : « le client a acheté le modèle standard » ;

M : « le client a acheté le modèle miniature » ;

A₁ : « le client a choisi l'abonnement d'une heure » ;

A₂ : « le client a choisi l'abonnement 2 h 30 ».

On note $p(E)$ la probabilité d'un événement E.

Les résultats seront donnés sous forme décimale avec 3 chiffres après la virgule.

1°) Déterminer $p(S)$, $p(M)$, $p(A_1)$.

2°) a) Parmi les clients qui ont acquis le modèle standard, 32 % ont pris l'abonnement A_1 . Traduire cette donnée en terme de probabilité.

b) En déduire la probabilité d'avoir acquis le modèle standard et d'avoir opté pour l'abonnement A_1 .

c) Justifier que la probabilité d'avoir choisi le modèle miniature et l'abonnement A_1 est égale à 0,288.

Dans cet énoncé l'expérience aléatoire est clairement précisée « Un client est pris au hasard dans l'échantillon ». Ces clients sont repérés par deux caractères statistiques ayant chacun deux modalités formant une partition de l'ensemble des clients. On peut donc résumer cet énoncé par un tableau de contingence. D'autre part les valeurs données sont des effectifs et il n'y apparaît pas de manière évidente de probabilités conditionnelles. Toutes les conditions semblent réunies pour s'aider d'un tableau, ce qui n'exclut pas la formalisation de la solution :

Tableau d'effectifs :

téléphone abonnement	téléphone		total par téléphone
	standard	miniature	
une heure			960
2 h 30			1 040
total par abonnement	1 200	800	2 000

duquel on déduit le tableau de fréquences ou de probabilités suivant :

événements	S	M	
A_1			0,48
A_2			0,52
	0,6	0,4	1

1°) D'après l'expérience aléatoire chaque client a la même probabilité d'être choisi, la probabilité de chaque événement le concernant sera donc égale à la fréquence observée dans l'échantillon, d'où les résultats écrits dans ce tableau.

2°) a) La donnée 32 % qui correspond au nombre décimal 0,32 est la probabilité conditionnelle qu'un client ait pris l'abonnement A_1 sachant qu'il est choisi parmi ceux qui ont acquis le modèle standard. On note : $p_S(A_1) = 0,32$.

b) La question posée revient à calculer $p(S \cap A_1)$, or on sait que :

$$p(S \cap A_1) = p(S) \times p_S(A_1) = 0,6 \times 0,32 = 0,192.$$

c) La question posée revient à calculer $p(M \cap A_1)$ et comme S et M forment une partition de Ω , on peut écrire :

$$p(S \cap A_1) + p(M \cap A_1) = p(A_1) \text{ d'où } p(M \cap A_1) = p(A_1) - p(S \cap A_1) = 0,48 - 0,192 = 0,288.$$

Cette solution est celle attendue par l'auteur de l'énoncé, mais l'élève qui a une conception cardinaliste très prégnante de la probabilité évitera de répondre à la question 2°) a), qui d'après lui ne sert à rien pour répondre aux autres questions qu'il résoudra alors, selon sa façon de concevoir la probabilité :

$$2^\circ) \text{ b) } 32 \% \text{ des } 1\,200 \text{ clients qui ont acquis le modèle standard représentent } 1\,200 \times \frac{32}{100} = 384$$

$$\text{clients qui ont pris l'abonnement } A_1, \text{ donc } p(S \cap A_1) = \frac{384}{2000} = 0,192.$$

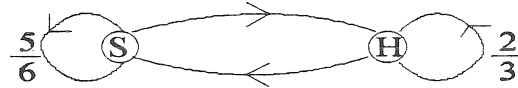
Cet élève écrira d'ailleurs cet effectif 384 dans la première case du tableau des effectifs et pourra ainsi répondre à la question 2°) c), en écrivant :

$$p(M \cap A_1) = \frac{960 - 384}{2\,000} = \frac{576}{2\,000} = 0,288.$$

Exercice 2

Exercice inspiré par A. ENGEL dans [3].

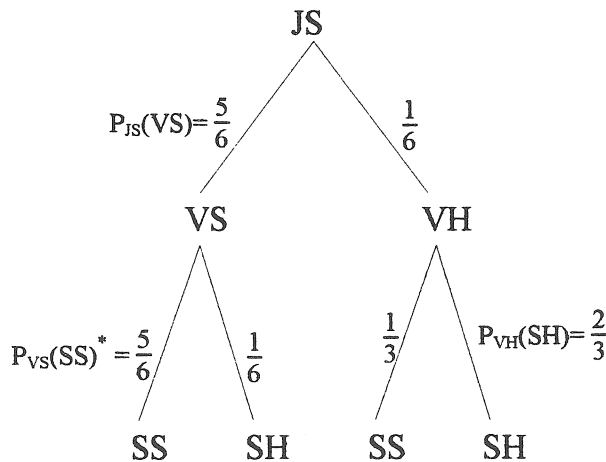
La figure suivante indique avec certitude(!) le comportement du temps en Probatatie, où on n'envisage que les possibilités : temps sec et temps humide.



Une étude météorologique a permis d'établir qu'en ce pays, s'il fait sec (S) aujourd'hui alors il fera encore sec demain avec la probabilité $\frac{5}{6}$ et s'il fait humide (H) aujourd'hui alors il fera encore humide demain avec la probabilité $\frac{2}{3}$. Aujourd'hui c'est jeudi et il fait sec, quelle est la probabilité d'avoir :

- 1°) samedi sec ?
- 2°) dimanche humide ?
- 3°) au moins une journée de pluie avant dimanche ?

Les données $\frac{5}{6}$ et $\frac{2}{3}$ de cet énoncé doivent être comprises comme des probabilités conditionnelles, ce qui oriente vers la construction d'un arbre probabiliste pour aider. Ce qui tient lieu d'événement certain c'est JS qui représente l'événement « jeudi sec ». De la même manière, on note VS et VH respectivement « vendredi sec » et « vendredi humide », etc. On obtient donc l'arbre probabiliste suivant :

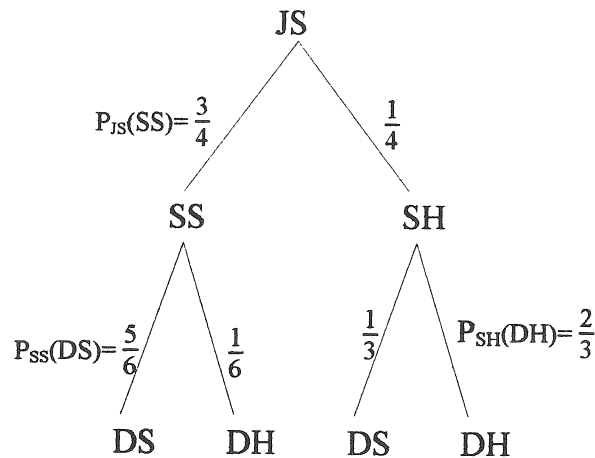


* Il faudrait normalement noter $P_{JS \text{ et } VS}(SS)$ mais d'après l'énoncé, la probabilité d'un événement ne dépend que de l'événement précédent, ce qui permet heureusement d'alléger la notation.

On peut donc répondre à la première question qui se traduit par :

$$P_{JS}(SS) = P_{JS}(VS \cap SS) + P_{JS}(VH \cap SS) = \frac{5}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{1}{3} = \frac{27}{36} = \frac{3}{4}$$

Pour répondre à la seconde question, on peut prolonger l'arbre précédent en ajoutant des branches, mais on peut auparavant le simplifier grâce à ce dernier résultat, ainsi on obtient :



On a donc : $P_{JS}(DH) = P_{JS}(SS \cap DH) + P_{JS}(SH \cap DH) = \frac{3}{4} \times \frac{1}{6} + \frac{1}{4} \times \frac{2}{3} = \frac{7}{24} \approx 0,29$.

La troisième question revient à calculer $P_{JS}(VH \cup SH)$:

On sait que $P_{JS}(VH \cup SH) = P_{JS}(VH) + P_{JS}(SH) - P_{JS}(VH \cap SH) = \frac{1}{6} + \frac{1}{4} - P_{JS}(VH) \times P_{JS \cap VH}(SH)$

d'où $P_{JS}(VH \cup SH) = \frac{1}{6} + \frac{1}{4} - P_{JS}(VH) \times P_{VH}(SH) = \frac{1}{6} + \frac{1}{4} - \frac{1}{6} \times \frac{2}{3} = \frac{6+9-4}{36} = \frac{11}{36} \approx 0,31$.

Si on avait choisi de noter P' la probabilité P_{JS} , les écritures des formules auraient été plus simples, en particulier la dernière s'écrirait alors :

$P'(VH \cup SH) = P'(VH) + P'(SH) - P'(VH \cap SH) = P'(VH) + P'(SH) - P'(VH) \times P'_{VH}(SH)$.

Pour répondre à cette troisième question, on aurait pu aussi s'intéresser à la variable aléatoire X du nombre de jours de pluie avant dimanche. La troisième question revient alors à calculer $P(X \geq 1) = 1 - P(X = 0) = 1 - P_{JS}(VS \cap SS)$

d'où $P(X \geq 1) = 1 - P_{JS}(VS) \times P_{JS \cap VS}(SS) = 1 - P_{JS}(VS) \times P_{VS}(SS) = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}$.

Peut-on poser la question ? :

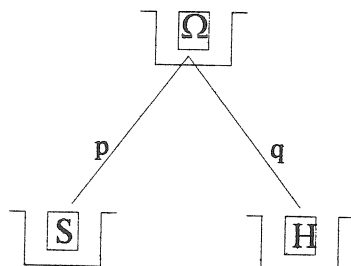
Quelle est la probabilité que le mardi précédent ait été humide ?

Certainement pas à nos élèves car on est dans la même situation qu'au début de l'exercice les jeunes et le chômage p.17. Pour répondre à la question posée, il faut, par exemple, se donner la probabilité p qu'un jour choisi au hasard dans ce pays soit sec d'où $q = 1 - p$ la probabilité que ce jour soit humide. On peut schématiser cette situation aléatoire par le modèle d'urnes suivant où toutes les boules sont de nature S ou H :

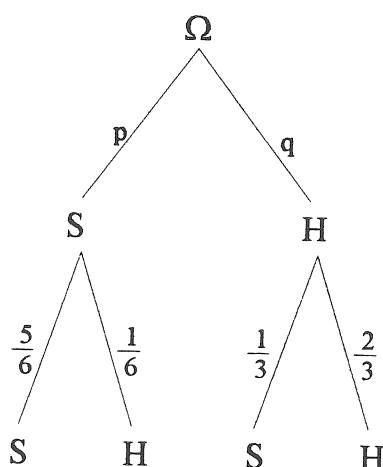
1°) Dans l'urne $\boxed{\Omega}$, les boules S sont en proportion p et on tire une boule au hasard.

2°) Si on tire une boule S alors on choisit ensuite une boule de l'urne $\boxed{S}$ où les boules S sont en proportion $\frac{5}{6}$, sinon on choisit dans $\boxed{H}$ où les boules H sont en proportion $\frac{2}{3}$.

3°) On replace la boule tirée dans son urne et on recommence ce protocole à partir du 2°).



De ce modèle d'urnes, on déduit la représentation en arbre probabiliste suivante :



Est-ce que les données de cet exercice sont suffisantes pour déterminer les probabilités p ou q ?

Notons P_n la probabilité que le $n^{\text{ième}}$ jour soit sec, on a :

$$P_1 = p$$

$$P_2 = p \times \frac{5}{6} + q \times \frac{1}{3} = p \times \frac{5}{6} + (1 - p) \times \frac{1}{3} = \frac{1}{2}p + \frac{1}{3} = \frac{1}{2}P_1 + \frac{1}{3}$$

Supposons que pour un entier $n \geq 2$, on ait : $P_n = \frac{1}{2}P_{n-1} + \frac{1}{3}$, alors $P_{n+1} = P_n \times \frac{5}{6} + (1 - P_n) \times \frac{1}{3}$,

soit : $P_{n+1} = \frac{1}{2}P_n + \frac{1}{3}$. Ainsi pour tout n de $\mathbb{N}^*$, on a : $P_{n+1} = \frac{1}{2}P_n + \frac{1}{3}$; c'est à dire que les différentes probabilités que les jours successifs soient secs forment une suite arithmético-géométrique de raison inférieure à 1, dont la limite l vérifie $\frac{1}{2}l + \frac{1}{3} = l$ soit $l = \frac{2}{3}$.

Si on considère les jours en nombre très grand avant et après ce Jeudi sec, on peut conclure que, quelle que soit la proportion observée p des jours secs en Probatatie, la probabilité qu'un jour lointain soit sec est très proche de $\frac{2}{3}$. On choisira donc $P(\text{MS}) = \frac{2}{3}$. Ce que confirment les calculs suivants des valeurs successives de P_n dans les cas où $p = 0,1$ puis $p = 0,9$:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
$p = 0,1$	0,383	0,525	0,596	0,631	0,649	0,658	0,662	0,664	0,666	0,666
$p = 0,9$	0,783	0,725	0,696	0,681	0,673	0,670	0,668	0,667	0,667	0,666

On peut alors apporter une réponse numérique à la question posée en utilisant la formule de Bayes :

Notons MS l'événement le mardi précédent a été sec, on a alors :

$$P_{\text{JS}}(\text{MS}) = \frac{P(\text{MS}) \times P_{\text{MS}}(\text{JS})}{P(\text{MS}) \times P_{\text{MS}}(\text{JS}) + P(\text{MH}) \times P_{\text{MH}}(\text{JS})}$$

$P_{\text{MS}}(\text{JS})$ est en fait la probabilité que 2 jours plus tard il fasse encore sec sachant qu'aujourd'hui il fait sec, c'est donc la même valeur que $P_{\text{JS}}(\text{SS})$ déjà calculée : $\frac{3}{4}$.

$P_{\text{MH}}(\text{JS})$ est la somme de deux probabilités conjointes, l'une utilisant l'éventualité Mercredi sec et l'autre Mercredi humide, on trouve : $\frac{1}{3} \times \frac{5}{6} + \frac{2}{3} \times \frac{1}{3} = \frac{9}{18} = \frac{1}{2}$.

$$\text{Ainsi } P_{\text{JS}}(\text{MS}) = \frac{\frac{2}{3} \times \frac{3}{4}}{\frac{2}{3} \times \frac{3}{4} + \frac{1}{3} \times \frac{1}{2}} = \frac{6}{6+2} = \frac{6}{8} = \frac{3}{4}.$$

On constate que la probabilité qu'il ait fait sec 2 jours auparavant est la même que la probabilité qu'il fasse sec 2 jours plus tard.

Exercice 3

Exercice commun à tous les candidats du Baccalauréat S., sujet national, juin 1998.

Dans tout l'exercice, A et B étant deux événements, $P(A)$ désigne la probabilité de A ; $P(B/A)$ la probabilité de B sachant que A est réalisé.

1. *Le nombre de clients se présentant en cinq minutes dans une station-service est une variable aléatoire X dont on donne la loi de probabilité :*

$$p_i = P(X = i)$$

i	0	1	2
p_i	0,1	0,5	0,4

- a. Définir et représenter graphiquement la fonction de répartition de X .
 - b. Calculer l'espérance mathématique de X .
2. *Dans cette station-service, la probabilité qu'un client achète de l'essence est 0,7 ; celle qu'il achète du gazole est 0,3. Son choix est indépendant de celui des autres clients. On considère les événements suivants :*
- C_1 : « en cinq minutes, un seul client se présente » ;
 C_2 : « en cinq minutes, deux clients se présentent » ;
 E : « en cinq minutes, un seul client achète de l'essence » ;
- a. Calculer $P(C_1 \cap E)$.
 - b. Montrer que $P(E/C_2) = 0,42$ et calculer $P(C_2 \cap E)$.
 - c. En déduire la probabilité qu'en cinq minutes un seul client achète de l'essence.
3. *Soit Y la variable aléatoire égale au nombre de clients achetant de l'essence en cinq minutes ; déterminer la loi de probabilité de Y .*

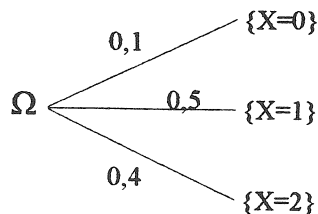
Dans cet exercice, la solution attendue par l'auteur ne semble pas inclure l'utilisation de schémas, néanmoins nous allons voir comment un arbre peut aider à la compréhension de la situation aléatoire²⁸ de cet énoncé. Nous allons constater comment **la construction rigoureuse de l'arbre probabiliste va nous obliger à approfondir et à expliquer en détail les hypothèses du modèle probabiliste choisi pour représenter cette situation.** Cette construction rend inévitable l'explicitation de ce modèle probabiliste, dont il ne faudrait pas, dans l'esprit des programmes actuels, pousser le formalisme trop loin.

²⁸ Bien sûr, un client ne choisit pas au hasard le type de carburant pour son véhicule, ici c'est son arrivée dans une station-service qui est considérée comme aléatoire.

En classe cependant, dans un objectif de formation, il serait bon de regarder avec les élèves, quels univers peuvent convenir à cette situation concrète. Il y en a qui sont "hors programme" comme ceux formés d'un produit cartésien, mais on peut s'intéresser à l'univers $\Omega = \{0, e, g, (e,e), (e,g), (g,e), (g,g)\}$ où ces issues, lors d'une tranche d'observation de 5 minutes, correspondent à :

- 0 pour aucun client ne se présente ;
- e pour un client se présente et achète de l'essence ;
- g pour un client se présente et achète du gazole ;
- (e,g) pour deux clients se présentent et achète dans l'ordre de l'essence puis du gazole ;
- etc.

La question 1. étant une question de cours intéressons-nous aux autres questions, après avoir placés les renseignements donnés dans l'énoncé sur un arbre probabiliste d'origine Ω .



On note E_i l'événement de Ω correspondant à l'événement familier :

« en cinq minutes, le $i^{\text{ème}}$ client achète de l'essence » ;

G_i pour « en cinq minutes, le $i^{\text{ème}}$ client achète du gazole » et

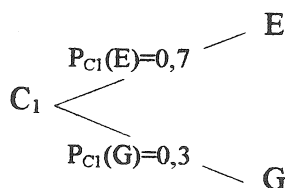
G pour « en cinq minutes, un seul client achète du gazole »²⁹

Pour la question 2. a., on admet que la phrase « Dans cette station-service, la probabilité qu'un client achète de l'essence est 0,7 ; celle qu'il achète du gazole est 0,3 » se traduit dans le modèle probabiliste (Ω, P) par : $P(E_1) = 0,7$ et $P(G_1) = 0,3$.

D'autre part, d'après la description de l'événement familier représenté par E dans le modèle, on a :

$P_{C_1}(E) = P(E_1)$; de même pour G : $P_{C_1}(G) = P(G_1)$.

Ainsi on obtient l'arbre d'origine C_1 suivant :



La réponse à la question posée est donc :

$$P(C_1 \cap E) = P(C_1) \times P_{C_1}(E) = P(\{X=1\}) \times P_{C_1}(E) = 0,5 \times 0,7 = 0,35$$

²⁹ En tant que partie de Ω , ces événements sont respectivement : $E = \{e, (e,g), (g,e)\}$; $E_1 = \{e, (e,g), (e,e)\}$; $E_2 = \{(g,e), (e,e)\}$ et $G = \{g, (e,g), (g,e)\}$; $G_1 = \{g, (g,e), (g,g)\}$; $G_2 = \{(e,g), (g,g)\}$

Pour la question 2.b., c'est un arbre d'origine C_2 qu'il faut considérer.

Si on note P' la probabilité conditionnelle sachant C_2 alors la phrase « Son choix est indépendant de celui des autres clients. » se traduit par :

1°) Comme pour toute épreuve de Bernoulli répétée :

$$P'_{E_1}(E_2) = P'(E_1) \text{ et } P'_{E_1}(G_2) = P'(G_1)$$

$$P'_{G_1}(E_2) = P'(E_1) \text{ et } P'_{G_1}(G_2) = P'(G_1)$$

2°) De plus, la notion d'indépendance utilisée par l'énoncé conduit à considérer que :

$$P'(E_1) = P(E_1) = 0,7 \text{ et } P'(G_1) = P(G_1) = 0,3.$$

D'où l'arbre suivant où Y est la variable aléatoire définie dans l'énoncé :

	événement-feuille de l'univers C_2 représenté ³⁰	probabilité P_{C_2} ou P' de cet événement
	$E_1 \cap E_2 = \{Y=2\}$	0,49
	$E_1 \cap G_2$	$\left. \begin{array}{l} \{0,21 \\ + \\ 0,21 \end{array} \right\} = E = \{Y=1\}$
	$G_1 \cap E_2$	
	$G_1 \cap G_2 = \{Y=0\}$	0,09

Les réponses aux questions posées sont alors :

Dans ce nouvel espace probabilisé, $(C_2, P_{C_2}) = (C_2, P')$, l'événement E est la réunion des événements disjoints $E_1 \cap G_2$ et $G_1 \cap E_2$ d'où :

$$P_{C_2}(E) = P'(E) = P'(E_1) \times P'_{E_1}(G_2) + P'(G_1) \times P'_{G_1}(E_2) = 0,7 \times 0,3 + 0,3 \times 0,7 = 0,42.$$

$$\text{Et } P(C_2 \cap E) = P(C_2) \times P_{C_2}(E) = P(\{X=2\}) \times 0,42 = 0,4 \times 0,42 = 0,168$$

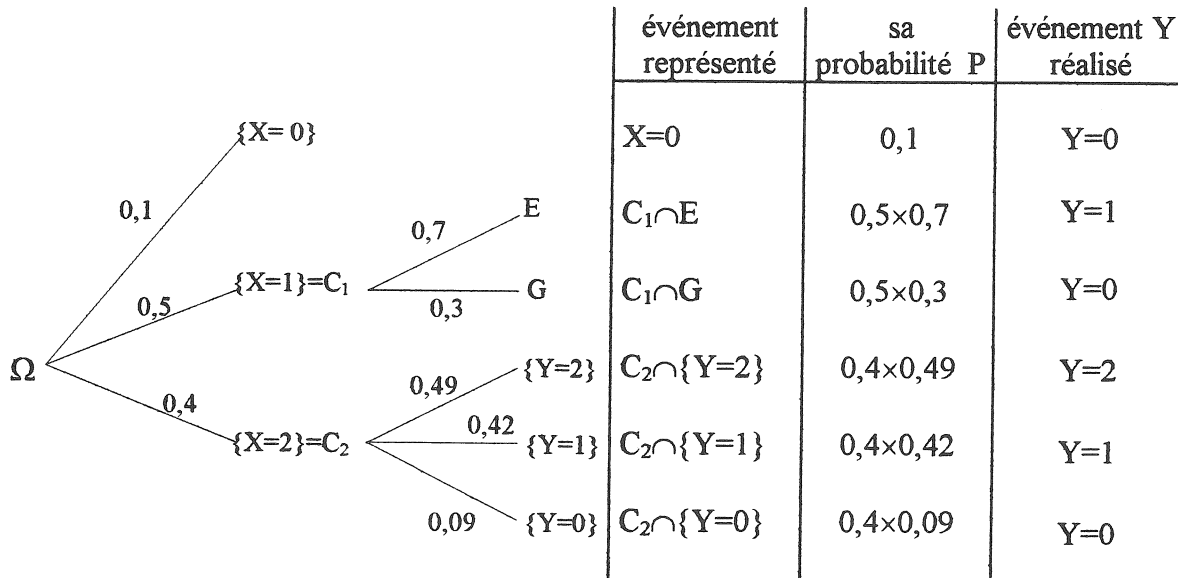
Pour la question 2.c., on écrit :

Dans l'espace probabilisé (Ω, P) , l'événement E est la réunion des événements disjoints $C_1 \cap E$ et $C_2 \cap E$ d'où $P(E) = P(C_1 \cap E) + P(C_2 \cap E) = 0,35 + 0,168 = 0,518$.

UNIVERSITÉ de LYON
BIBLIOTHEQUE
Université Claude Bernard - LYON I
du 11 Novembre 1983
VILLEURBANNE Cedex

³⁰ ou encore, trace dans C_2 de l'événement de Ω .

Pour répondre à la question 3., il est nécessaire d'avoir une vue d'ensemble de cette situation. Celle-ci peut nous être donnée par l'arbre probabiliste suivant. Cet arbre s'obtient en raccordant les deux arbres précédents d'origine C_1 et C_2 à l'arbre d'origine Ω de la première question :



Il ne reste plus qu'à justifier pourquoi l'événement $\{Y=2\}$ est représenté par une seule feuille de cet arbre, d'où la réponse : $P(Y=2) = P(C_2 \cap \{Y=2\}) = 0,4 \times 0,49 = 0,196$

Pour $\{Y=1\}$, représenté par deux feuilles de cet arbre, la réponse a été demandée à la question 2. c. : $P(Y=1) = P(E) = 0,518$.

Et pour $\{Y=0\}$, on peut le considérer comme événement complémentaire de $\{Y=1\} \cup \{Y=2\}$, d'où la réponse : $P(Y=0) = 1 - 0,196 - 0,518 = 0,286$.

On peut aussi contrôler ce dernier résultat en faisant la somme des probabilités des trois événements-feuilles qui le constituent.

CONCLUSION

Pour résoudre un exercice de probabilité, il y a en général une des trois approches décrites précédemment qui est mieux appropriée que les deux autres. L'enseignant doit donc surmonter sa propre approche, qui est souvent inconsciente, pour proposer à ses élèves des exercices où interviennent équitablement ces différentes façons. C'est lorsqu'on a dépassé les différentes façons de concevoir la probabilité qu'on est arrivé au concept mathématique de la probabilité.

Si on choisit une représentation congrue avec un énoncé, la solution sera plus simple et ne nécessitera pas une résolution comportant des inconnues ou des questions intermédiaires surabondantes. Parmi ces représentations, les **arbres probabilistes** sont des supports très pertinents pour organiser des données, s'approprier certains énoncés et élaborer les raisonnements à mettre en oeuvre, **en particulier en probabilités conditionnelles**. Ils fournissent une aide très efficace à la compréhension des concepts et à l'assimilation des théorèmes ou formules de probabilités.

B. PARZYSZ dans [7] et dans [1] p.233 à 238 valide l'arbre en tant qu'outil de résolution de problème, en explicite des règles de traitement et de conversion, il en indique la portée et les limites. Dans un langage didacticien, il précise que *« l'articulation cadre/conversion et registre/traitement permet en particulier de distinguer ce qui est propre au domaine mathématique considéré et ce qui peut éventuellement être utilisé dans d'autres contextes (statistique, dénombrement) »*. Il ne faudrait certes pas abuser des arbres car des élèves pourraient croire que l'objet de l'apprentissage est le mode d'emploi des arbres, le concept de probabilité conditionnelle ne serait alors pas acquis. Il y aurait évitement de l'obstacle qui ne serait donc pas surmonté comme le signale M. HENRY dans [5].

Il apparaît donc souhaitable que les élèves sachent utiliser d'autres supports visuels comme les **tableaux à double entrée** (diagrammes de Carroll) qui ont l'avantage de ne pas être séquentiels. Il semble également souhaitable qu'ils voient l'utilisation des arbres, pondérés ou non, dans d'autres situations comme par exemple les dénombrements.

Concernant ces schémas et ces outils en Probabilités, d'importantes questions pédagogiques se posent :

Notre enseignement peut-il se satisfaire du choix par l'élève de la représentation-outil la mieux adaptée à la situation aléatoire rencontrée ? ou encore ;

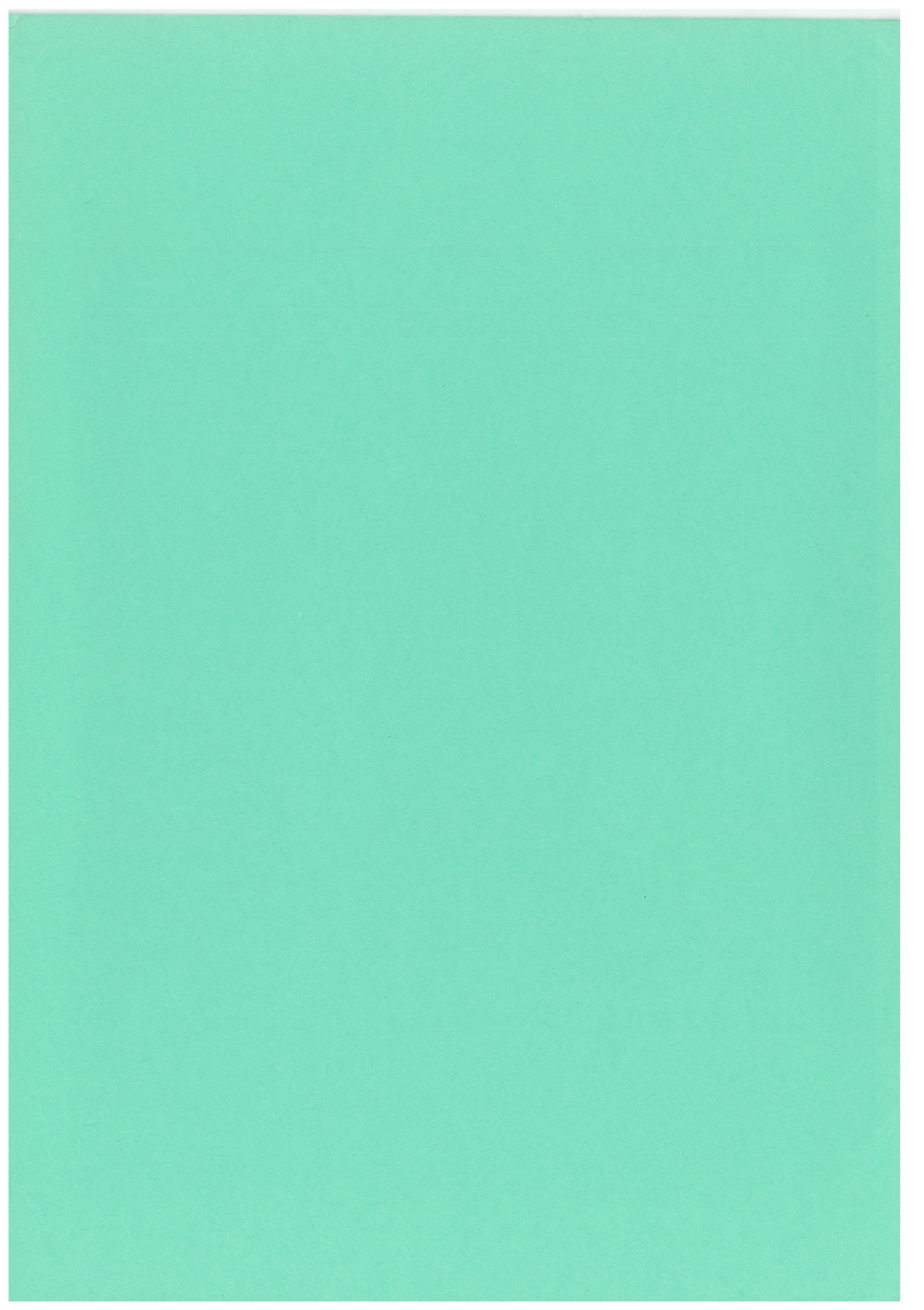
Doit-on enseigner systématiquement ces outils et leurs fonctionnements, et entraîner les élèves à passer de l'un aux autres comme le proposent C. DUPUIS et S. ROUSSET-BERT dans [2] ?

Remarquons que ces questions se posent également en Analyse avec les représentations graphiques et d'autres schémas comme les tableaux de variation.

Il semble que l'enseignant doit principalement garder présent à l'esprit que toutes ces représentations, aussi pertinentes et formatrices soient elles, doivent être passagères pour peu à peu **laisser place à l'écriture formelle des calculs probabilistes**, relevant d'une compréhension en profondeur des outils probabilistes de base, gage de leur efficacité dans les conditions les plus générales et seule garantie de la validité mathématique des résultats obtenus.

BIBLIOGRAPHIE

- [1] COMMISSION inter-IREM « STATISTIQUE ET PROBABILITÉS » (Juin 1997) : *Enseigner les probabilités au lycée.*
- [2] Claire DUPUIS et Suzette ROUSSET-BERT (Janvier 1996) : *Arbres et tableaux de probabilité* dans Repères-IREM n° 22
- [3] Arthur ENGEL (1975) : *L'enseignement des probabilités et de la statistique* (adapté de l'allemand), éditions CEDIC; (épuisé), réédité aux éditions ALÉAS.
- [4] Jean-Pierre GRANGÉ (Mai 1996) : *Probabilités conditionnelles et indépendance.* -IREM de BESANÇON
- [5] Michel HENRY (Oct 1994) : *L'enseignement des probabilités : perspectives historiques, épistémologiques et didactiques.* -IREM de BESANÇON.
- [6] Michel HENRY (1994) : *L'enseignement du calcul des probabilités dans le second degré.* - dans Repères-IREM n° 14.
- [7] Bernard PARZYSZ (1993) : *Des statistiques aux probabilités : exploitons les arbres.* -dans Repères-IREM n° 10.
- [8] André TOTOHASINA (1994) : *Introduction du concept de probabilité conditionnelle : avantages et inconvénients de l'arborescence.* -dans Repères-IREM n° 15.



I.R.E.M. de Franche-Comté
UFR des Sciences et Techniques
16, route de Gray, La Bouloie
F-25030 BESANÇON cedex
Tél. : 03.81.66.61.92 - Fax : 03.81.66.61.99
Courrier électronique : iremfc@math.univ-fcomte.fr
<http://pegase.univ-fcomte.fr/CTU/IREM/lieux.htm#Besançon>

TITRE : ARBRES ET PROBABILITES

AUTEURS : Jean-Pierre GRANGE

PUBLIC CONCERNE : PROFESSEURS DE
MATHEMATIQUES ENSEIGNANT EN LYCEE OU
FORMATEURS INTERVENANT EN FORMATION
INITIALE

DATE : Mars 1999

MOTS CLES : Probabilités, représentations, modélisations,
arbres pondérés, arbres probabilistes, règles d'arbres,
tableaux de contingence.

RESUME : Cette brochure explicite trois façons différentes
de concevoir la probabilité : cardinaliste, fréquentiste et
formelle. On y voit leur influence sur l'enseignement, sur
la formulation des exercices ainsi que sur les
représentations des probabilités.

Des tableaux à double entrée (diagrammes de Carroll) sont
utilisés comme représentation-outil pour résoudre des
exercices et les arbres pondérés ou arbres probabilistes y ont
une place de choix dans les applications et les exercices. Des
règles de construction et d'utilisation de ces arbres sont
également précisées.

Pour ceux qui aiment les chiffres, (oh pardon! les nombres)
cette brochure contient 21 arbres et 4 tableaux.

Format A4 - Nombre de pages : 43 - Poids : 130 g
Prix : 30 F

IREM DE BESANÇON

Dépot Légal: 99/117

Numéro ISBN : 2.909963.24.1