Faites vos jeux

Introduction au calcul des probabilités

rédigée par D. Degen et M. Krysinska,
en collaboration avec B. Jadin, dans le cadre de séminaires du GEM

Ce qui est improbable n'est pas impossible.

W. Shakespeare

1 Estimation de chances

Dans ce chapitre, on présente des premières expériences aléatoires, des exemples d'un événement impossible, d'un événement certain et des moyens pour comparer les chances de réalisation de différents événements.

1.1 Introduction

L'approche du concept de probabilité développée dans ce chapitre s'appuie sur ses deux facettes.

La première facette est liée à la structure physique du mécanisme du hasard et se base sur la loi expérimentale selon laquelle les fréquences relatives d'un événement aléatoire ont tendance à se stabiliser. Le calcul des fréquences induit le calcul des probabilités.

La seconde facette est liée à notre connaissance du phénomène étudié, au degré de nos convictions ou croyances. Le calcul des chances ou distribution des probabilités est souvent basé sur l'équiprobabilité liée aux symétries d'une expérience.

Même si la théorie des probabilités ne se développe qu'au XVII^{e} siècle, la prise de conscience des phénomènes dus au hasard sous ces deux facettes est présente dès l'Antiquité. Citons, pour exemple, l'astragale, l'ancêtre du dé. C'est un des os du pied d'un animal (mouton, chèvre, cerf,...) qui, une fois lancé, retombe sur l'une de ses faces.

D'une part, les expériences avec d'anciens astragales montrent que les fréquences des différentes faces ne sont pas les mêmes. Cela amène les joueurs à observer les résultats et à garder les fréquences en mémoire. Les fréquences expérimentales sont donc aussi anciennes que le jeu d'astragale.

D'autre part, on a trouvé sur d'anciens sites quelques astragales parfaitement symétriques et polis en forme de cube. Cela laisse supposer que ceux qui les ont conçus avaient associé la symétrie de l'objet à une même fréquence pour chacune des faces.

La rencontre entre le calcul des chances et le calcul des fréquences a permis de fonder le calcul des probabilités au milieu du XVIIe siècle : Pascal (1623-1662) et Fermat (1601-1665), Cette rencontre a été réalisée à l'occasion du problème des *partis*, c'est-à-dire des partages de gains au cours d'un jeu de hasard.

1.2 Pile ou face

À combien de chances estimez-vous la possibilité d'obtenir « Pile » lors d'un lancer d'une pièce ? Pourquoi ?

Voici des réponses possibles :

- La chance d'obtenir « Pile » est égale à une chance sur deux car il y a deux résultats possibles « Pile » et « Face ».
- On peut avoir un troisième résultat «Tranche ».
- Il y a une chance sur deux car les faces de la pièce sont symétriques.
- Lorsqu'on lance la pièce un grand nombre de fois, on peut s'attendre à obtenir « pile » dans, à peu près, 50% des cas.

La première réponse n'est pas correcte car on peut avoir une fausse pièce où les chances d'obtenir « pile » ou « face » ne sont pas les mêmes.

Le troisième résultat possible « Tranche » se produit rarement car il est lié avec l'équilibre fragile de la pièce dans cette position. On dira dans ce cas que ce résultat est possible mais improbable.

Les deux dernières réponses font apparaître deux facettes du concept de probabilité : l'une s'appuie sur notre confiance en la symétrie de la pièce et l'autre fait appel à la prévision des résultats lors de nombreux lancements.

1.3 Lancer une punaise

Vous avez devant vous une punaise. Quelles sont les chances qu'elle retombe « tête en bas » ? Faites des expériences pour évaluer ces chances.

On peut lancer une punaise 10 fois, 20 fois, 30 fois, … On note après chaque série de 10 lancers, la fréquence de la première issue possible de l'expérience, à savoir « tête en bas »,. On obtient aussi celle de la seconde issue possible de l'expérience, à savoir « pointe en bas ». Cela devrait nous permettre d'y voir plus clair quant à la propension de cette punaise particulière de retomber sur l'une ou sur l'autre position.

Pour accélérer l'expérience, on peut lancer plusieurs punaises toutes de même gabarit, en même temps. On n'observe plus tout à fait la même chose, puisqu'il s'agit ici de la propension d'un type de punaise, et non plus d'une punaise donnée, à retomber sur la tête ou sur la pointe.

Nous avons répété 10 fois l'expérience avec 50 punaises. Nous avons calculé la fréquence de l'issue « tête en bas », autrement dit le rapport du nombre de « tête en bas » au nombre total de punaises, après le premier lancer, après deux lancers, après trois lancers, et ainsi de suite. Les résultats du calcul sont présentés dans le tableau 1.

« Tête en bas	Fréquences
Série 1	0,500
Séries 1 et 2	0,500
Séries 1, 2 et 3	0,493
Séries 1, 2, 3 et 4	0,470
Séries 1, 2, 3, 4 et 5	0,476
Séries 1, 2, 3, 4, 5 et 6	0,497
Séries 1, 2, 3, 4, 5, 6 et 7	0,500
Séries 1, 2, 3, 4, 5, 6, 7 et 8	0,493
Séries 1, 2, 3, 4, 5, 6, 7, 8 et 9	0,489
Séries 1, 2, 3, 4, 5, 6, 7, 8, 9 et 10	0,484

Tabl. 1

On observe que les fréquences tendent à se stabiliser (Fig.1). Cette stabilisation sera, de surcroît, observée chaque fois qu'on va refaire l'expérience. Par exemple, en considérant les lancers de tous les élèves d'une classe (tous ayant utilisé le même type de punaises et s'étant mis d'accord sur un processus de lancement), on obtiendra une autre fréquence qu'on pourra utiliser pour évaluer la probabilité de « tête en bas ».

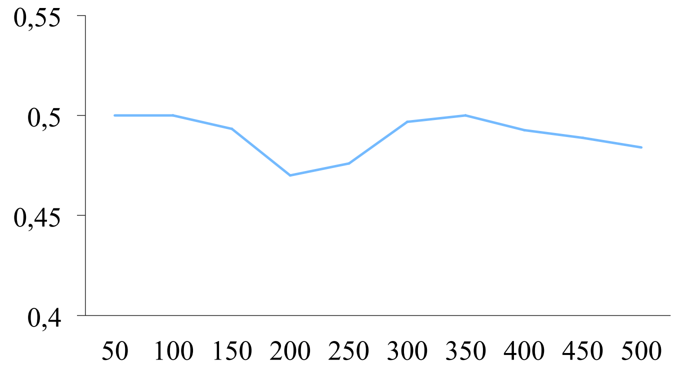


Fig. 1 : Evolution de fréquences « tête en bas » en fonction du nombre des lancers de punaises

Ces expériences montrent qu'il n'est pas possible de déterminer exactement la chance de l'issue « tête en bas ». Nous pouvons très bien décider que cette chance est égale à soit à 0,48 , soit à 0,49 ou encore à 0,495….. ou à tout autre nombre proche de ceux observés lors des expérimentations. Lorsque nous prenons la décision de choisir un nombre, par exemple 0,48 comme probabilité de l'issue « tête en bas », ce nombre correspond à la mesure du degré de confiance en la réalisation de cet événement. Cela signifie que, dans une longue série de lancers d'une même punaise, cet événement sera réalisé à peu près dans 48% des lancers.

Les deux exemples traités ci-dessus, celui de la pièce et celui de la punaise, illustrent comment on fait le choix d'un **modèle de probabilités**. Dans le cas d'une pièce de monnaie, nous avons choisi le **modèle théorique** d'équiprobabilité en exprimant ainsi notre confiance en la symétrie de la pièce. En cas de doute, nous aurions pu choisir un **modèle expérimental** en s'appuyant sur les fréquences observées, comme dans l'exemple de la punaise.

1.4 Fréquence de naissances des filles et des garçons

La naissance d'un enfant, c'est un peu comme jouer à pile ou face, si ce n'est qu'on ne répète pas l'expérience autant de fois qu'on veut. Certains pensent qu'il y a autant de chance d'avoir une fille que d'avoir un garçon, autrement dit que la probabilité de naître fille est la même que celle de naître garçon, c'est-à-dire 0,5. Les relevés statistiques réalisés en Belgique entre 1982 et 1992 (source INS) présentés dans le tableau 2 confirment-ils cette opinion ?

Année	Nombre de naissances		
	Total	Garçons	Filles
1982	120 241	61 831	58 410
1983	117 145	60 209	56 936
1984	115 651	59 331	56 320
1985	114 092	58 572	55 520
1986	117 114	60 493	56 621
1987	117 354	60 390	56 964
1988	119 779	61 520	58 259
1989	120 904	61 948	58 956
1990	123 776	63 421	60 335
1991	125 924	64 586	61 338
1992	124 774	63 883	60 891

Tabl. 2

Intéressons-nous à la fréquence du nombre de naissances (tableau 3), d'une année à l'autre.

Années	Nombre de naissances				
	Total	Garçons	Filles	%Garçons	%Filles
1982	120 241	61 831	58 410	51 ,42	48 ,58
1983	117 145	60 209	56 936	51 ,40	48 ,60
1984	115 651	59 331	56 320	51 ,30	48 ,70
1985	114 092	58 572	55 520	51 ,34	48 ,66
1986	117 114	60 493	56621	51 ,65	48 ,35
1987	117 354	60 390	56 964	51 ,46	48 ,54
1988	119 779	61 520	58 259	51 ,36	48 ,64
1989	120 904	61 948	58 956	51 ,24	48 ,76
1990	123 776	63 421	60 335	51 ,24	48 ,46
1991	125 924	64 586	61 338	51 ,29	48 ,71
1992	124 774	63 883	60 891	51 ,20	48 ,80

Tabl. 3

Au départ, on observe les fluctuations. Après avoir cumulé les fréquences sur 1, 2, …, 11 ans (Tabl. 4), celles-ci se stabilisent. On peut estimer les chances d'avoir une fille à 0,4865 c'est-à-dire à 48,65 % et un garçon à 51,35 %.= 0,5135.
Remarquons que 0,4865 + 0,5135 = 1. Ceci correspond au fait que les naissances des filles et celle des garçons représentent ensemble 100% des naissances et mène à la règle suivante :
<u>La somme des probabilités d'événements complémentaires est toujours égale à 1.</u>

Comme dans le cas de la punaise, nous pouvons choisir un modèle expérimental de probabilités, par exemple 0,486 pour la naissance d'une fille, comme nous pouvons aussi bien décider de choisir un autre nombre proche de ceux qui sont observés dans le tableau. Nous pouvons aussi décider d'attribuer le nombre 0,5 comme probabilité de la naissance d'une fille. Dans ce cas, nous faisons le choix du modèle théorique d'équiprobabilité.

1.5 Lancer un dé

On lance un dé cubique. Lequel des événements suivants est probable, certain ou improbable ?
Obtenir un nombre pair.
Obtenir un nombre composé de deux chiffres.
Obtenir un nombre plus petit que 10.

L'événement « obtenir un nombre pair » est un événement probable ; la valeur de la probabilité y associée est un nombre strictement compris entre 0 et 1.
L'événement « obtenir un nombre composé de deux chiffres » est un événement impossible. Nous attribuons la valeur 0 à la probabilité de cet événement.
L'événement « obtenir un nombre plus petit que 10 » est un événement certain. Nous attribuons la valeur 1 à la probabilité d'un tel événement.

1.6 Choisir une urne

Dans une urne, il y a 4 boules noires et 6 boules blanches. Dans une autre urne, il y a 8 boules noires et 12 boules blanches. Laquelle des deux urnes choisir pour avoir le plus de chance de tirer une boule noire ?

Certains peuvent penser que les chances de tirer une boule noire sont plus grandes dans l'urne où les boules noires sont plus nombreuses. Or, dans les deux urnes, le rapport entre le nombre de boules noires et le nombre de boules rouges est le même, donc le mécanisme aléatoire est en quelque sorte, le même. Cela permet d'attribuer la probabilité 0,4 à l'événement « tirer une boule noire » et la probabilité 0,6 à l'événement « tirer une boule rouge ». Implicitement, on s'appuie ici sur l'hypothèse que n'importe quelle boule de la première urne est tirée avec la même probabilité égale à 0,1=1/10 et que n'importe quelle boule de la deuxième urne est tirée avec la même probabilité égale à 0,05 =1/20. D'où l'on peut considérer que

$$0,4 = 4/10 = 8/20 \quad \text{et} \quad 0,6 = 6/10 = 12/20$$

Cette question permet de dégager la règle suivante :
<u>Lorsqu'on fait le choix du modèle théorique d'équiprobabilité, la probabilité d'un événement est égale au rapport du nombre de résultats (cas ou issues) favorables à la réalisation de l'événement considéré au nombre total des résultats (cas) possibles.</u>

1.7 Problème de Galilée [Engel, 1975]

Le prince de Toscane demanda un jour à Galilée : " Pourquoi, lorsque l'on jette trois dés, obtient-on plus souvent la somme 10 que la somme 9 bien que ces deux sommes soient obtenues de six façons différentes ? "

À première vue, les événements « obtenir la somme 9 » et « obtenir la somme 10 » auraient pu être considérés comme équiprobables car chacun d'eux est réalisé de 6 façons différentes. En effet, la somme «9» peut s'obtenir comme

1+2+6, 1+3+5, 2+3+4, 2+2+5, 1+4+4, 3+3+3

et la somme «10» peut aussi s'obtenir comme

1+3+6, 2+3+5, 1+4+5, 2+2+6, 3+3+4, 2+4+4.

Mais les observations faites par le prince de Toscane issues de son expérience de joueur contredisent l'hypothèse de l'équiprobabilité. Où se cache donc l'erreur ? Lorsqu'on suppose que les dés sont parfaitement équilibrés, alors toutes les issues d'un lancer de trois dès sont équiprobables. Plusieurs de ces issues donnent une même somme. Par exemple, chaque lancer de trois dés qui réalise l'évènement « 6+2+1 » est composé de six issues dites « favorables » :

$1^{\text{ère}}$ dé	2^{e} dé	3^{e} dé
6	2	1
6	1	2
2	6	1
2	1	6
1	6	2
1	2	6

Tabl. 4

Par contre, il n'y a qu'une seule issue favorable à l'évènement « 3+3+3 » et trois issues favorables à l'évènement comme 6+2 +2 :

$1^{\text{ère}}$ dé	2^{e} dé	3^{e} dé
6	2	2
2	6	2
2	2	6

Tabl. 5

Au total, il y a

- 25 issues équiprobables et favorables à « 9 » : six issues pour 6+2+1, six issues pour 5+3+1, six issues pour 4+2+3, trois issues pour 5+2+2, trois issues pour 4+1+4, une issue pour 3+3+3 ;
- 27 issues équiprobables et favorables à « 10 » : six issues pour 6+3+1, six issues pour 5+4+1, six issues pour 2+3+5, trois issues pour 6+2+2, trois issues pour 4+3+3, trois issues pour 4+4+2.

Ainsi, on voit qu'il y a plus d'issues équiprobables réalisant la somme « 10 » que d'issues équiprobables réalisant la somme « 9 », ce qui se traduit par la plus grand fréquence de la somme « 10 » par rapport à ma somme « 9 » lors d'un grand nombre d' expériences répétées.

Calculons à présent les probabilités d'obtenir la somme « 10 » et la somme « 9 ». Pour cela, nous devons connaître le nombre total d'issues possibles. La technique du calcul qui appartient au domaine de l'analyse combinatoire, sera la suivante. Il y a 6 issues différentes du jet du premier dé ; à chaque issue ainsi obtenue, correspondent 6 issues possibles du jet du second dé ; à chaque issue du jet de deux premiers dés correspondent 6 issues du jet du troisième dé, ce qui fait au total $6 \times 6 \times 6 = 216$ issues possibles.

La probabilité d'obtenir « 9 » comme somme est alors égale à $\frac{25}{216} = 0{,}115740\ldots$ et celle d'obtenir « 10 » est égale à $\frac{27}{216} = 0{,}125$. Cela signifie qu'on peut s'attendre à ce que, sur 1000 jets de trois dés, on obtient « 9 » presque 116 fois et « 10 » à peu près 125 fois. La différence n'est pas grande mais pour un joueur invétéré, elle peut devenir importante.

Synthèse : Estimation de chances

Précisons quelques mots de vocabulaire :
- **probabilité** est d'origine latine « probabilis » et signifie probable ou vraisemblable ;
- **aléatoire** est d'origine latine « alea » et signifie jeu de dés ;
- **chance** est d'origine latine « cadere » et signifie choir (chute de dé) ;
- **hasard** est d'origine arabe «az zahr » et signifie jeu de dés ;

Une **expérience aléatoire** est une action ou un processus qui peut avoir plusieurs résultats mais celui qui sera réalisé ne peut être prédit avec certitude.

Voici quelques exemples d'expériences aléatoires :

- lancer un dé et noter le nombre de points apparus sur la face supérieure ; les résultats possibles ou issues sont « obtenir 1 »,« obtenir 2 », « obtenir 3 », « obtenir 4 », « obtenir 5 » ou « obtenir 6 » ;
- tirer une boule d'une urne dans laquelle se trouvent des boules blanches et noires et noter la couleur de la boule tirée ; les résultats possibles ou issues sont « tirer une blanche » ou « tirer une noire » ;
- enregistrer une naissance et observer le sexe de l'enfant ; les résultats possibles ou issues sont « le nouveau-né est un garçon » ou « le nouveau-né est une fille ».

Un **événement** relatif à une expérience aléatoire est composé d'un ou de plusieurs des résultats ou issues. Par exemple, dans le jet de trois dés, l'évènement « avoir la somme 9 » est composé des résultats ou issues non équiprobables comme 1+2+6, 1+3+5, 2+3+4, 2+2+5, 1+4+4, 3+3+3 et il est composé de 25 résultats ou issues équiprobables sous la forme des triplets de nombres de 1 à 6 dont la somme est égale à 9.

Un **événement certain** est toujours réalisé, tandis que un **événement impossible** ne se réalise jamais. Par exemple, lors du lancer d'un dé cubique, « Obtenir l'un des nombres 1, 2, 3, 4, 5, 6 » est un événement certain ; «Obtenir 7 » est un événement impossible.

L'**événement contraire** ou **complémentaire** d'un événement donné se réalise lorsque ce dernier ne se réalise pas.
Par exemple, dans un lancer de dé, « Obtenir un nombre impair » est l'événement contraire de l'événement « Obtenir un nombre pair ». On notera par *non A* l'événement contraire de l'événement *A* associé à une même expérience aléatoire.

Des **événements incompatibles** relatifs à une même expérience aléatoire sont des événements qui ne peuvent pas se réaliser en même temps.
Par exemple, dans un lancer de dé, « Obtenir un nombre strictement plus petit que 3» et « Obtenir un nombre strictement plus grand que 4 » sont des événements incompatibles. Des événements contraires sont toujours incompatibles tandis que des événements incompatibles ne sont pas forcément contraires comme le montre l'exemple utilisé ci-dessus.

Une expérience aléatoire est entièrement définie par une liste des résultats ou évènements ou issues possibles appelée **système complet d'événements** ou **catégorie d'épreuves** ou encore **univers,** qui doivent être deux à deux incompatibles et tels qu'un seul d'entre eux est réalisé à l'issue d'une expérience.

Les questions 1.1, 1.2 et 1.3 ont permis de formuler une toute première approche de la notion de **probabilité.**

La **probabilité d'un événement A** est un nombre compris entre 0 et 1 qui correspond

- soit au degré de confiance, de conviction que nous avons dans le fait que cet événement se produise (par exemple 0,5 pour obtenir « Pile » ou « Face » lors du lancer d'une pièce lorsque nous sommes convaincus de la symétrie de la pièce). Cette approche correspond à un **modèle théorique de probabilités**.

- soit à la fréquence $f_n = \frac{nombre\ de\ réalisations\ de\ A}{n}$. avec laquelle cet événement s'est produit dans le passé lors d'une longue série d'expériences (par exemple l'événement « Tête en bas » lors du lancer d'une punaise). Cette approche correspond à un **modèle expérimental de probabilités**, il s'appuie sur le phénomène de stabilisation des fréquences d'un événement.

Le choix d'un modèle de probabilités pour une expérience aléatoire permet de faire des prévisions quant aux fréquences d'événements liés à cette expérience: une probabilité associée à un événement est une sorte de fréquence théorique c'est-à-dire une fréquence attendue d'un événement lors d'un grand nombre de répétitions de l'expérience. Par exemple, la probabilité 1/6 associée à l'événement « obtenir 5 » lors du lancer d'un dé, signifie que lors d'un très grand nombre de lancers, on peut s'attendre théoriquement (donc sans faire l'expérience) à obtenir 5 une fois sur six.

La probabilité permet de prévoir approximativement la fréquence avec laquelle un événement considéré se produira au cours d'une longue série de répétitions d'une même expérience aléatoire.

Parmi les modèles théoriques, on privilégie le **modèle d'équiprobabilité** lorsqu'on a la conviction que tous les résultats d'une expérience aléatoire ont la même chance de se réaliser. Ce modèle a été mis en place dans la question 1.4 relative au choix d'une urne.

Lorsque tous les résultats sont équiprobables, la probabilité d'un événement réalisé par plusieurs de ces résultats est égale au rapport du nombre de résultats qui réalisent cet événement (nombre de cas favorables) au nombre total des résultats possibles (nombre de cas possibles)**.**

Nous avons utilisé cette propriété dans les questions 1.6 et 1.7.

Première règle du calcul des probabilités : la somme des probabilités d'événements complémentaires est toujours égale à 1.

Nous avons mis en place cette règle dans la question 1.4.

Lien entre probabilités et fréquences empiriques

Dans la question 1.2, « lancers des punaises », nous avons observé la stabilisation des fréquences cumulées avec l'augmentation du nombre des lancements des punaises. La stabilisation observée est prévue par le théorème de la théorie des probabilités qu'on appelle la loi des grands nombres. Voici en substance son contenu. On considère une expérience aléatoire dans laquelle le hasard intervient pour déterminer une issue. Soit *A* un événement, résultat possible de cette expérience, constitué par certaines issues de probabilité théorique p. On répète cette expérience un nombre n de fois et on calcule la fréquence f_n des réalisations de A. La loi des grands nombres affirme que, plus n est grand, plus on a de chance que l'écart entre f_n et p soit plus petit que n'importe quel nombre positif donné. Cette loi permet de calculer par exemple qu'avec $n > 1000$ il y a plus de 95 chances sur 100 que la différence $|f_n - p|$ soit inférieure à 0,03.

Cette loi justifie donc le choix de la fréquence de l'événement « tête en bas » dans un grand nombre de répétitions de l'expérience « lancer la punaise », ou le choix d'un nombre proche de cette fréquence, comme probabilité de l'évènement « tête en bas ».

2 Probabilités composées

Dans ce chapitre, on modélise à l'aide de diagrammes en arbre des expériences composées de plusieurs étapes et on met en place les règles d'utilisation de ces diagrammes. Ces diagrammes servent ensuite comme outil de résolution de divers problèmes.

2.1 Tables de mortalité

Tout comme le sexe d'un enfant, la durée de vie est également imprévisible. Beaucoup de contrats comme ceux d'assurance vie, de rente viagère, par exemple, reposent sur l'estimation de celle-ci. D'où la nécessité d'estimer selon l'âge du bénéficiaire le nombre d'années qu'il a encore à vivre.

Les tables de mortalité du tableau 6 reprennent pour un million de naissances et selon le sexe, le nombre de survivants à chaque âge. Elles sont fournies par l'INS en 2003 et élaborées à partir d'un très grand nombre d'observations.

a) Qui a le plus de chances d'atteindre 85 ans : un garçon de 16 ans ou un homme de 75 ans ? Cette impression est-elle confirmée par les données du tableau 7 ?

b) À partir des données du tableau 6, calculez la probabilité qu'un nouveau-né garçon atteigne l'âge de 85 ans, celle qu'un nouveau-né garçon atteigne l'âge de 40 ans et celle qu'un homme de 40 ans atteigne l'âge de 85 ans. Quel est le lien entre ces trois probabilités ? Calculez la probabilité qu'un nouveau-né fille atteigne l'âge de 85 ans.

c) À partir des données du tableau 6 et en s'appuyant sur le modèle expérimental des naissances des filles et des garçons, calculez la probabilité qu'un nouveau-né soit un garçon et atteigne l'âge de 85 ans et celle qu'un nouveau-né soit une fille et atteigne l'âge de 85 ans.

d) À partir des données du tableau 6 et en s'appuyant sur le modèle expérimental des naissances des filles et des garçons, déterminez la probabilité qu'un nouveau-né atteigne l'âge de 85 ans.

e) Quelle est la probabilité qu'un acheteur verse encore une rente 10 ans après la signature d'un contrat de vente viagère, lorsque le vendeur est un couple dont l'homme est âgé de 75 ans et la femme est âgée de 70 ans?

x	Homme	Femme	*x*	Homme	Femme	*x*	Homme	Femme
0	1 000 000	1 000 000						
1	999 415	999 664	41	964 210	982 167	81	498 950	670 315
2	998 827	999 328	42	962 327	981 327	82	466 949	641 506
3	998 237	998 990	43	960 312	980 427	83	434 106	610 817
4	997 644	998 653	44	958 152	979 461	84	400 631	578 299
5	997 047	998 314	45	955 834	978 421	85	366 772	544 045
6	996 447	997 974	46	953 341	977 299	86	332 810	508 203
7	995 843	997 634	47	950 658	976 084	87	299 058	470 977
8	995 234	997 292	48	947 766	974 768	88	265 855	432 633
9	994 621	996 949	49	944 645	973 337	89	233 556	393 505
10	994 002	996 605	50	941 273	971 779	90	202 525	353 988
11	993 377	996 258	51	937 628	970 079	91	173 113	314 540
12	992 745	995 910	52	933 683	968 222	92	145 653	275 668
13	992 106	995 560	53	929 411	966 190	93	120 434	237 916
14	991 459	995 208	54	924 782	963 962	94	97 692	201 840
15	990 802	994 852	55	919 763	961 517	95	77 591	167 983
16	990 136	994 494	56	914 321	958 831	96	60 210	136 845

17	989 459	994 133	57	908 417	955 877	97	45 543	108 845
18	988 769	993 768	58	902 011	952 625	98	33 492	84 294
19	988 067	993 398	59	895 060	949 044	99	33 492	84 294
20	987 349	993 024	60	887 519	945 096	100	16 452	46 069
21	986 616	992 645	61	879 339	940 744	101	10 917	32 273
22	985 864	992 260	62	870 468	935 942	102	6 951	21 688
23	985 093	991 868	63	860 853	930 645	103	4 228	13 915
24	984 300	991 469	64	850 437	924 800	104	2 446	4 874
25	983 483	991 062	65	839161	918 351	105	1 339	4 874
26	982 640	990 646	66	826 964	911 238	106	690	2 627
27	981 768	990 220	67	813 786	903 393	107	332	1 317
28	980 864	989 782	68	799 564	894 747	108	149	609
29	979 925	989 331	69	784 236	885 223	109	61	258
30	978 947	988 866	70	767 741	874 740	110	23	99
31	977 927	988 386	71	750 022	863 213	111	8	34
32	976 860	987 887	72	731 027	850 551	112	2	10
33	975 742	987 369	73	710 708	836 663	113	1	3
34	974 567	986 828	74	689 026	821 452	114	0	1
35	973 330	986 263	75	665 955	804 822	115	0	0
36	972 024	985 671	76	641 480	786 678	116	0	0
37	970 644	985 047	77	615 606	766 928	117	0	0
38	969 181	984 389	78	588 355	745 485	118	0	0
39	967 627	983 693	79	559 774	722 274	119	0	0
40	965 973	982 954	80	529 938	697 231	120	0	0

Tabl. 6

a) Les réponses divergent : pour les uns, c'est le garçon de 16 ans et pour d'autres c'est un homme de 75 ans qui a plus de chances d'atteindre 85 ans . C'est cette dernière réponse qui est correcte : cela peut paraître surprenant au premier abord, mais cela s'explique lorsqu'on se rend compte qu'un garçon de 16 ans doit d'abord survivre jusque 75 ans avant d'atteindre 85 ans. Parmi tous les garçons de 16 ans, une partie d'entre eux seulement atteindra l'âge de 75 ans.

Le calcul suivant confirme ce raisonnement. En effet, parmi les 990136 garçons de 16 ans, il y en a 366772 qui survivent jusqu'à 85 ans; ce qui veut dire que la proportion de garçons de 16 ans qui survivent jusqu'à 85 ans vaut $\frac{366772}{990136} = 0{,}370426 \ldots$, tandis que parmi les 665955 hommes de 75 ans, il y en a 366772 qui survivent jusqu'à 85 ans, d'où la proportion des hommes de 75 ans qui survivent jusqu'à 85 ans vaut $\frac{366772}{665955} = 0{,}550545 \ldots$

Puisque cette table de mortalité a été élaborée à partir d'un très grand nombre d'observations, on peut considérer que 0,370426 est la probabilité qu'un garçon de 16 ans atteigne l'âge de 85 ans et que $0{,}550545$ est la probabilité qu'un homme de 75 ans atteigne l'âge de 85 ans.

b) La probabilité qu'un nouveau-né garçon atteigne l'âge de 85 ans peut être obtenue de deux manières différentes :

- d'une part, les tables de mortalité nous indiquent que sur un million de garçons qui naissent, il n'en reste, 85 ans plus tard, que 366772. Dès lors la probabilité qu'un nouveau-né garçon atteigne l'âge de 85 ans peut être estimée par $\frac{366772}{1000000} = 0{,}36672 \ldots \approx 37\%$.

- d'autre part et de manière analogue, on peut estimer la probabilité qu'un nouveau-né garçon atteigne l'âge de 40 ans par $\frac{965973}{1000000} = 0{,}965973 \ldots \approx 96{,}6\%$ et celle qu'un homme de 40 ans atteigne l'âge de 85 ans par $\frac{366772}{965973} = 0{,}379691 \approx 38\%$

Ces trois fractions sont liées par l'égalité suivante : $\frac{965973}{1000000} \times \frac{366772}{965973} = \frac{366772}{1000000}$, ce qui est équivalent à la multiplication de deux probabilités : $0{,}965973 \times 0{,}379691 = 0{,}36672$.

En rapprochant ces deux modes de calcul, on illustre une règle du calcul des probabilités appelée **règle de multiplication.** La probabilité qu'un nouveau-né garçon atteigne l'âge de 85 ans est égale à la probabilité qu'un nouveau-né garçon atteigne l'âge de 40 ans multipliée par la probabilité qu'un homme de 40 ans atteigne l'âge de 85 ans.

Ces deux modes de calcul correspondent à deux expériences aléatoires différentes :

- la probabilité $0{,}366772$ correspond à l'expérience aléatoire simple qui consiste à tirer au sort un homme parmi un million d'hommes. On a deux résultats possibles: l'homme a survécu jusqu'à l'âge de 85 ans ou l'homme est décédé avant.

- le produit $0{,}965973 \times 0{,}379691$ correspond à une autre expérience aléatoire qui comprend deux étapes. À la première étape, on tire au sort un homme parmi un million d'hommes. On a deux résultats possibles ; l'homme a survécu jusqu'à l'âge de 40 ans ou il est décédé avant. À la deuxième étape, on tire au sort un homme parmi ceux qui ont survécu jusqu'à l'âge de 40 ans et l'on a deux résultats possibles : l'homme a survécu jusqu'à l'âge de 85 ans ou l'homme est décédé avant.

À chacune de ces expériences, on peut associer un **diagramme en arbre des possibles** qui reprend les résultats possibles de l'une et de l'autre.

Le diagramme en arbre correspondant à la première expérience est très simple (Fig.2) :

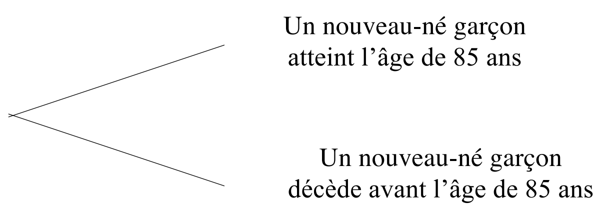


Fig.2 : Diagramme en arbre de l'expérience simple « un nouveau-né garçon atteint l'âge de 85 ans »

Le diagramme en arbre correspondant à la deuxième expérience se compose de deux étapes (Fig.3) :

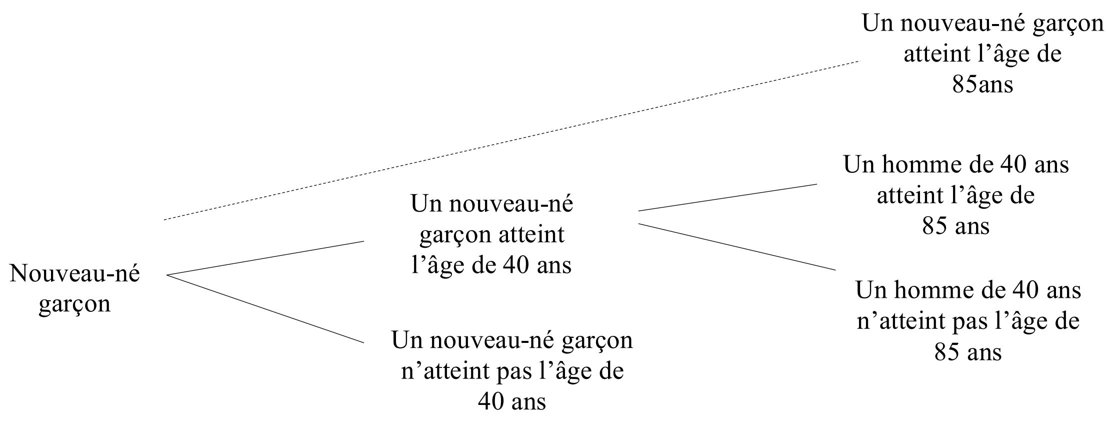


Fig.3 : Diagramme en arbre des possibles de l'expérience composée « un nouveau-né garçon atteint l'âge de 40 ans et un homme de 40 ans atteint l'age de 85 ans »

Ajoutons, dans les nœuds de cet arbre, les données numériques tirées du tableau 7 (Fig.4).

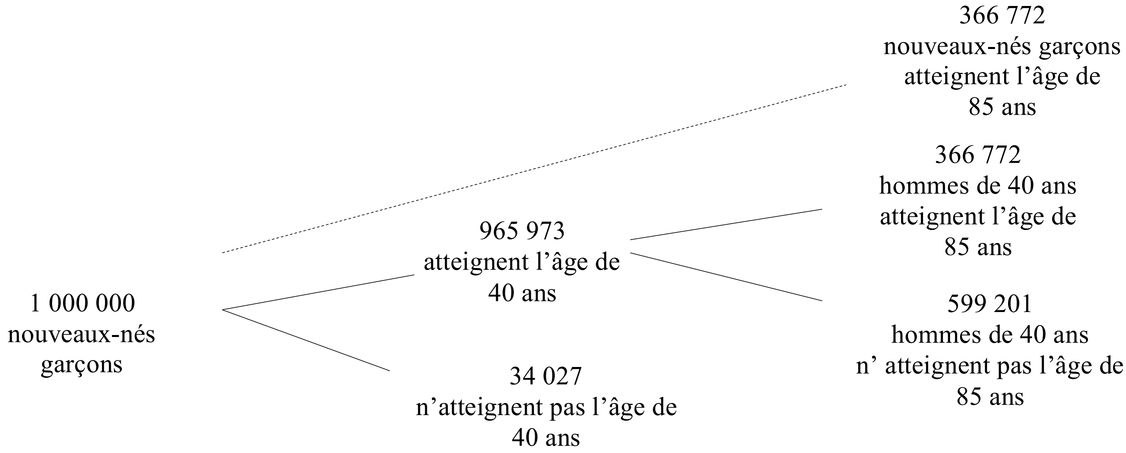


Fig.4 : Diagramme en arbre avec les données numériques de l'expérience composée - un nouveau-né garçon atteint l'âge de 40 ans et un homme de 40 ans atteint l'age de 85 ans

Le nombre d'un million de nouveaux-nés garçons est arbitraire. Quel que soit, au départ, le nombre *N* de nouveaux-nés garçons, nous retrouvons le même type de calcul pour déterminer les données numériques correspondant à *N* nouveaux-nés (Fig.5).

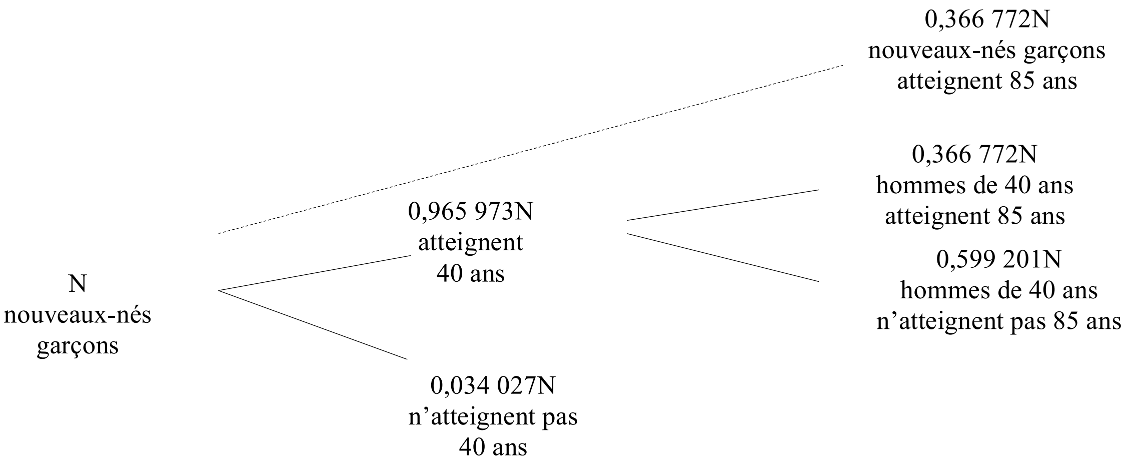


Fig.5 : Diagramme en arbre généralisé de l'expérience composée - un nouveau-né garçon atteint l'âge de 40 ans et un homme de 40 ans atteint ou pas l'age de 85 ans

Que nous utilisions l'arbre de la figure 4 ou l'arbre de la figure 5 pour calculer la probabilité des événements suivants, nous obtenons le même résultat. En effet,

- la probabilité qu'un nouveau-né garçon atteigne l'âge de 85 ans vaut

$$\frac{366772}{1000000} = \frac{0{,}366772N}{N} = 0{,}366772,$$

- la probabilité qu'un nouveau -né garçon atteigne l'âge de 40 ans vaut

$$\frac{965973}{1000000} = \frac{965973N}{N} = 0{,}965973,$$

- la probabilité qu'un nouveau-né garçon décède avant l'âge de 40 ans vaut

$$\frac{34027}{1000000} = \frac{0{,}34027N}{N} = 0{,}34027\ ;$$

- la probabilité qu'un homme de 40 ans atteigne l'âge de 85 ans vaut

$$\frac{366772}{965973} = \frac{0{,}366772N}{0{,}965973N} = 0{,}379691,$$

- la probabilité qu'un homme de 40 ans décède avant l'âge de 85 ans vaut

$$\frac{599201}{965973} = \frac{0{,}599201N}{0{,}965973N} = 0{,}620308,$$

Par convention, on décide de noter ces probabilités le long des branches (Fig.6). On obtient le **diagramme en arbre probabilisé** :

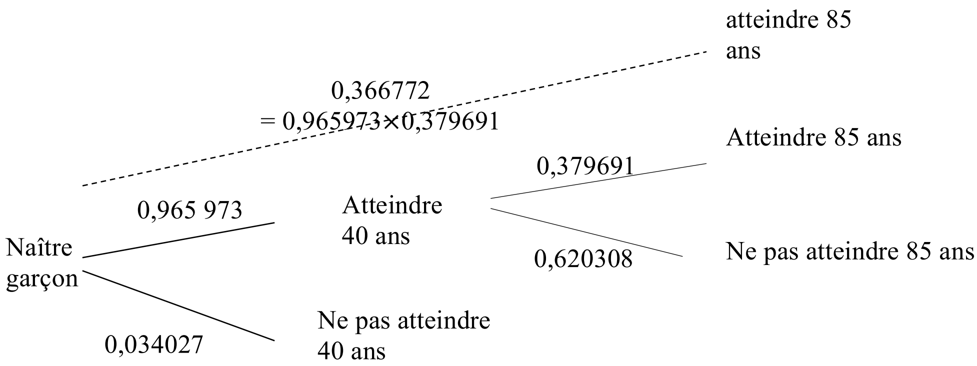


Fig.6 : Diagramme probabilisé de l'expérience composée - un nouveau-né garçon atteint l'âge de 40 ans et un homme de 40 ans atteint l'age de 85 ans

La **règle de multiplication** énoncée au début de cette section peut être formulée ainsi : *Lorsqu'on parcourt un chemin (branches consécutives) d'un arbre, la probabilité composée est obtenue en multipliant les probabilités écrites au-dessus de chacune des branches rencontrées.*

De manière analogue, on calcule que la probabilité qu'un nouveau-né fille atteigne l'âge de 85 ans est de 0,544045.

c) Pour répondre à la question posée, une manière de procéder est de déterminer le nombre des garçons et des filles atteignant l'âge de 85 ans pour un très grand nombre de nouveaux-nés, par exemple, un million. Si on admet que la probabilité de naître fille est de 48,65 %, il y a de fortes chances d'avoir sur un million de nouveaux-nés 486500 filles et 513500 garçons.
Parmi ces 513500 garçons, le nombre d'hommes qui atteignent l'âge de 85 ans est de 0,366772 × 513 500 = 188 337.
En procédant de la même façon pour les 486500 filles, on obtient 264678 femmes qui atteignent l'âge de 85 ans.
Ces calculs peuvent être organisés sous forme d'un diagramme (Fig.7) :

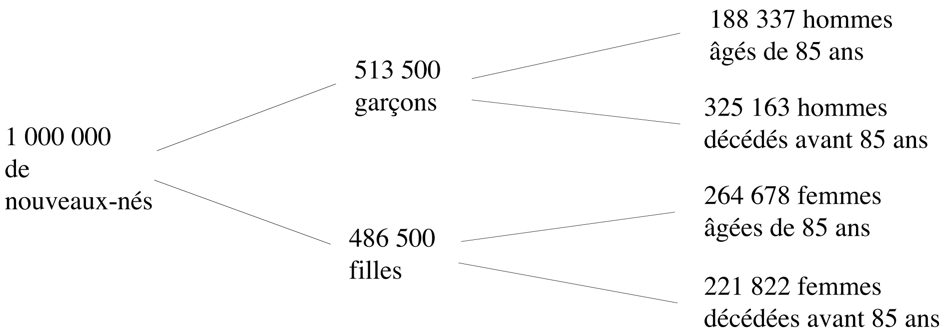


Fig.7 : Diagramme complété par les données numériques de l'expérience composée - un nouveau-né atteint l'âge de 85 ans

Mettons en place le diagramme en arbre probabilisé (Fig.8) :

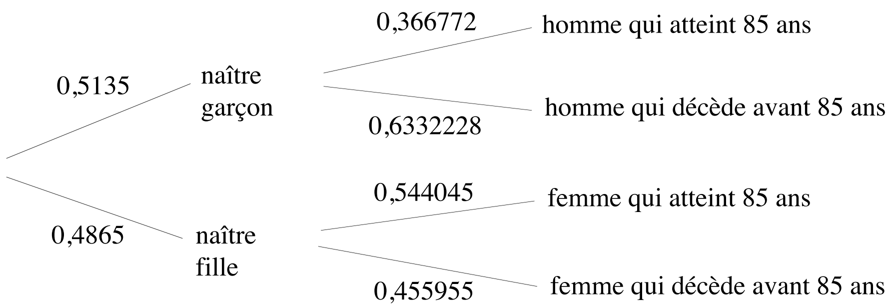


Fig.8 : Diagramme probabilisé de la survie d'un nouveau-né jusqu'à l'âge de 85 ans

La probabilité qu'un nouveau-né soit un garçon **et** atteigne l'âge de 85 ans vaut 0,188337 et est obtenue comme produit 0,513500 × 0,366772. Il s'agit d'une **probabilité composée**. De même, la probabilité qu'un nouveau-né soit une fille **et** atteigne l'âge de 85 ans vaut 0,264677 et obtenu comme produit 0,4865 ×0,544045.

d) Pour répondre à la question posée, une manière de procéder est de déterminer le nombre de survivants à l'âge de 85 ans pour un très grand nombre de nouveaux-nés, par exemple, un million. D'après les calculs faits en c), parmi ces survivants, il y a 188 337 hommes de 85 ans et 264678 femmes de 85 ans. Au total, sur un million de nouveaux-nés, on a donc de fortes chances de retrouver 188337 + 264678 = 453015 personnes de 85 ans. Ainsi, la probabilité qu'un nouveau-né vive jusqu'à l'âge de 85 ans est égale à $\frac{453015}{1000000} = 0{,}453015$.

Regardons de plus près la structure de ce calcul :

$$\begin{aligned}
0{,}453015 &= \frac{453015}{1000000} \\
&= \frac{188337+264678}{1000000} \\
&= 0{,}188337 + 0{,}264678 \\
&= 0{,}366772 \times 0{,}513500 + 0{,}544045 \times 0{,}486500
\end{aligned}$$

Il illustre la deuxième règle d'utilisation des arbres de probabilités appelée **règle d'addition** qui s'énonce comme suit :

Lorsqu'on doit parcourir plusieurs chemins d'un même arbre pour rencontrer tous les événements faisant partie d'un événement composé, la probabilité totale de celui-ci est égale à la somme des probabilités correspondant à chacun des chemins.

e) La question posée revient à chercher la probabilité qu'au moins un des deux conjoints soit encore en vie après 10 ans. Désignons par « H » l'événement « un homme âgé de 75 ans atteint l'âge de 85 ans » et par « non H » l'événement contraire « un homme âgé de 75 ans décède avant d'atteindre l'âge de 85 ans ». De même, désignons par « F » l'événement « une femme âgée de 70 ans atteint l'âge de 80 ans » et par « non F » l'événement contraire « une femme âgée de 70 ans décède avant d'atteindre l'âge de 80 ans ». La figure 9 présente deux diagrammes en arbre des possibles.

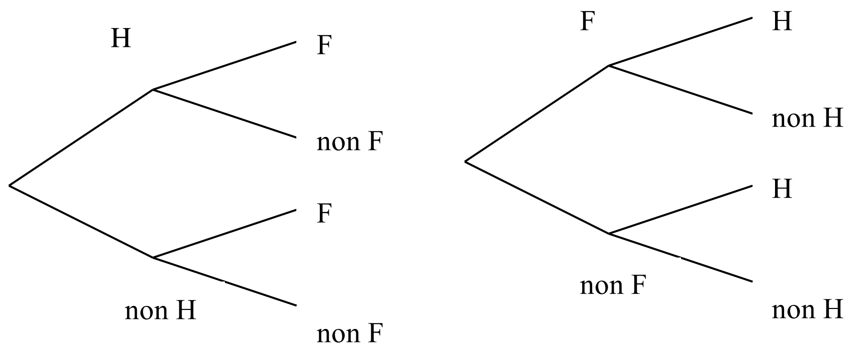
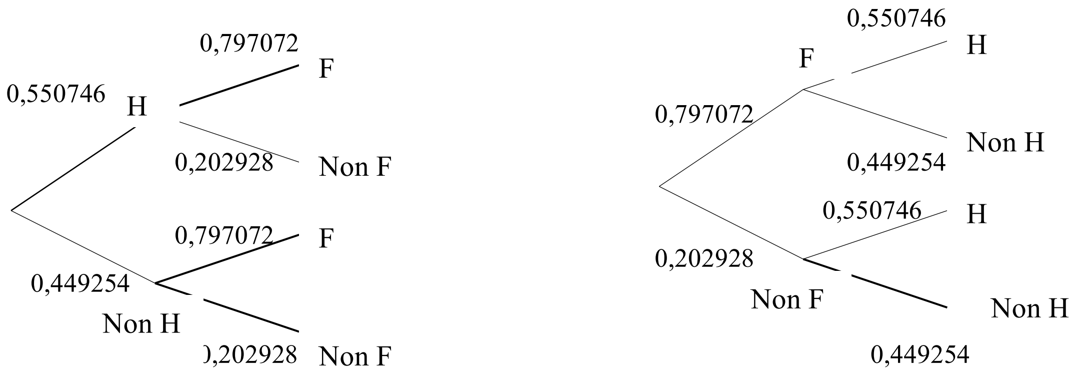


Fig. 9 : Diagrammes en arbre des possibles de la survie des deux conjoints

En faisant l'hypothèse que la durée de vie d'un des conjoints n'est pas influencée par celle de l'autre, les arbres partiels de la figure 10 porteront les mêmes probabilités, quelle que soit leur position dans l'arbre de la figure 9. La probabilité de l'événement « H » est égale à 0,550746 (question a)) et donc celle de « non H » sera égale à 1– 0,550746 =0,449254. A partir du tableau 7, on établie que la probabilité de « F » est égale à $\frac{697231}{874740} = 0,797072$ et celle de « non F » est égale à $1 - 0,797072 = 0,202928$. Ainsi, les diagrammes des possibles de la figure 9 peuvent être probabilisé (Fig.10)

Fig. 10 : Diagrammes en arbre partiels probabilisés de survie de chacun des conjoints

La figure 11 présente les diagrammes de la figure 9 complétée par les probabilités associées à chaque branche.



a) b)

Fig. 11 : Diagrammes en arbre complets et probabilisés de la survie du couple

Pour rencontrer l'événement « au moins un des deux conjoints est encore en vie après 10 ans », on doit retenir les trois premiers chemins des figures 11a) ou 11b). En appliquant les règles de multiplication et d'addition respectivement aux deux arbres de ces diagrammes, on obtient un même résultat :

$$0{,}550746 \times 0{,}797072 + 0{,}550746 \times 0{,}202928 + 0{,}449254 \times 0{,}797072 = 0{,}908834$$

$$0{,}797072 \times 0{,}550746 + 0{,}797972 \times 0{,}449254 + 0{,}202928 \times 0{,}550746 = 0{,}908834$$

Le quatrième chemin correspond au décès des deux conjoints ou, ce qui revient au même, à l'événement contraire de l'événement «l'acheteur ne paie plus de rente ». Cela signifie que la probabilité que nous cherchions aurait pu être calculée plus simplement comme suit :

$$1 - 0{,}440254 \times 0{,}202928 = 0{,}908834$$

2.2 Famille de quatre enfants

Dans les familles de quatre enfants que rencontre-t-on le plus souvent : deux filles et deux garçons ou trois enfants d'un même sexe ?

Certains pensent que 50% des familles de quatre enfants comprend deux filles et deux garçons, comme pour les chances d'avoir une fille ou d'avoir un garçon mais ne savent pas quel pourcentage attribué aux familles comprenant trois enfants d'un même sexe.

D'autres pensent à la probabilité 1/3 pour les deux types de famille, (car il y a des familles soit avec un garçon, soit avec deux garçons, soit avec trois garçons). D'autres encore pensent que cette probabilité vaut 1/5 (car il y a en plus des familles composées seulement de filles et des familles composées seulement de garçons.)

Ceux qui n'ont pas d'idées, se mettent à répertorier les familles de quatre enfants qu'ils connaissent ... Cette dernière réponse est loin d'être dépourvue d'intérêt. Elle suggère l'expérimentation. Mais, comment expérimenter? En effet, répéter l'expérience un grand nombre de fois s'avère difficile et l'on ne dispose pas de relevés statistiques. Il reste donc à simuler l'expérience avec une table de nombres aléatoires. Un logiciel ou une calculatrice programmable peuvent aussi générer des nombres aléatoires. La simulation montrerait probablement que les chances d'avoir trois enfants du même sexe est plus grande que celle d'avoir deux filles et deux garçons. Tentons d'expliquer cet avis.

Quel que soit le raisonnement utilisé, il faut tenir compte du choix d'un modèle de probabilités. Soit on considère que les probabilités de naître fille ou garçon sont égales Soit on considère qu'elles sont respectivement égales à 0,48 et à 0,52 pour respecter la réalité observée sur des relevés statistiques.

La figure 13 présente un diagramme en arbre des résultats possibles lors des naissances successives dans une famille de quatre enfants ; nous avons noté par « F » la naissance d'une fille et par « G » celle d'un garçon.

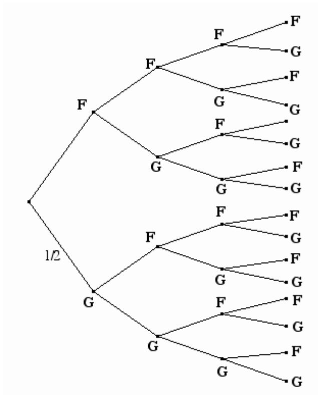


Fig. 12 : Diagramme en arbre des possibles de la composition des familles de quatre enfants

À partir de cet arbre, on peut répondre en termes de « chances » : il y a 6 chances sur 16 d'avoir deux filles et deux garçons et 8 chances sur 16 d'avoir trois enfants d'un même sexe. En répondant de cette façon, on suppose que les 16 résultats possibles ont la même chance de se réaliser. Cette hypothèse est-elle valable ? Pour le vérifier, transformons l'arbre des possibles en arbre probabilisé en affectant à chaque branche la probabilité de passer d'un nœud à l'autre. Insistons encore sur le fait que préalablement à la détermination des probabilités, le choix d'un modèle doit être fait.

En choisissant le modèle P (F) = P (G) = 1/2, on obtient le diagramme en arbre suivant (

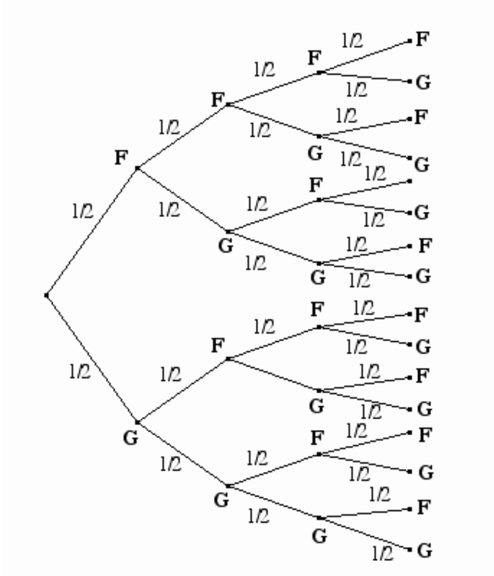


Fig.13 : Diagramme en arbre probabilisé de la composition des familles de quatre enfants

La probabilité d'un événement réalisé à l'extrémité de chaque chemin, par exemple, P (FFFG), est de $\frac{1}{16}$.

L'événement « 2F et 2G » est réalisé à l'extrémité de 6 chemins différents, chacun d'eux étant affecté d'une probabilité égale à $\frac{1}{16}$, on a

$$\text{P (2F et 2G)} = \frac{6}{16} = 0,375$$

De la même manière, on a

$$P\text{ (3F et 1G ou 3G et 1F)} = \frac{8}{16} = 0{,}5.$$

Ces deux derniers calculs s'appuient sur les règles d'addition et de multiplication formulées dans le chapitre 1.

Le choix du modèle probabiliste où P(F)= 0,4865 et P(G)= 0,5135, mène à des chemins non équiprobables et en procédant selon les mêmes règles de calcul, on obtient les résultats suivants :

$$P\text{ (2F-2G)} \cong 0{,}3745 \quad \text{et } P\text{ (3F-1G ou 3G-1F)} \cong 0{,}4999997$$

Ces résultats nous montrent qu'indépendamment du choix d'un des deux modèles probabilistes, les familles avec deux enfants de même sexes sont moins probables que les familles avec trois enfants de même sexe dont la probabilité est soit égale à 0,5, soit très proche de 0,5.

2.3 Problème du condamné à mort (inspiré par [Engel, 1975])

À un condamné à mort, on laisse une dernière chance. Il doit, les yeux bandés, choisir l'une des trois urnes de la figure 14 et tirer une boule de cette urne. Tirer une boule blanche lui sauve la vie.

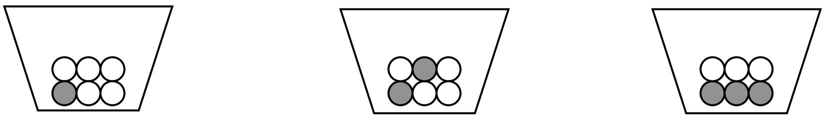


Fig.14 : Une répartition des boules blanches et noires dans les urnes

Quelles sont ses chances de survie ?
Si on laisse au condamné la possibilité de répartir comme il le désire les boules dans les urnes, peut-il augmenter ses chances de survie ?

a) L'expérience aléatoire consiste à choisir au hasard d'abord une urne et ensuite, à tirer une boule dans l'urne choisie. Cette expérience peut être schématisée à l'aide du diagramme en arbre suivant (Fig.15).

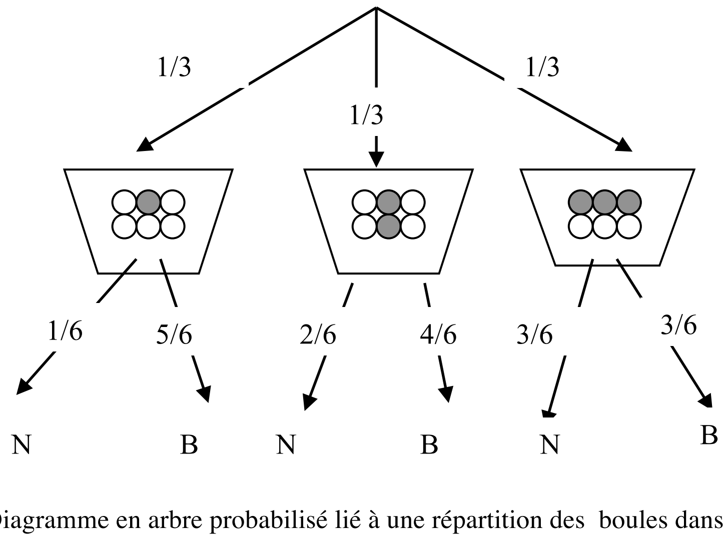


Fig. 15 : Diagramme en arbre probabilisé lié à une répartition des boules dans des urnes

Le calcul des chances de survie du condamné revient à calculer la probabilité de tirer une boule blanche $P(A)$, elle est égale à $P(B) = \frac{1}{3}\times\frac{5}{6} + \frac{1}{3}\times\frac{4}{6} + \frac{1}{3}\times\frac{3}{6} = \frac{12}{18} = \frac{2}{3}$.

b) On pense fréquemment qu'une autre répartition des boules ne « change rien ». En effet, quelle que soit la répartition, il y a toujours 12 boules blanches sur les 18 boules et donc le condamné a toujours 12 chances sur 18 de tirer une boule blanche.
Pour appuyer cette opinion, certains donnent l'exemple d'une autre répartition (Fig.16) qui mène à la même probabilité de survie $P(B) = \frac{1}{3}\times 1 + \frac{1}{3}\times 1 + \frac{1}{3}\times 0 = \frac{2}{3}$.

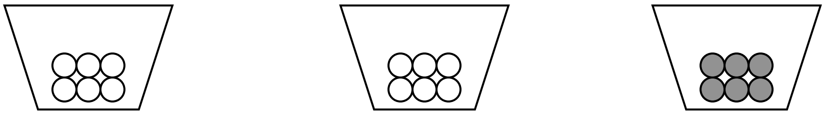


Fig. 16 : Une autre répartition des boules blanches et noires dans les urnes

Cet exemple permet à d'autres de comprendre qu'il ne sert à rien de mettre autant de boules blanches dans les deux premières urnes. En effet, pour être certain de tirer une boule blanche dans une de ces urnes, il suffit de mettre une seule boule blanche dans chacune d'elles. Ainsi dans la dernière urne, on aura six boules noires et dix boules blanches (Fig.17).

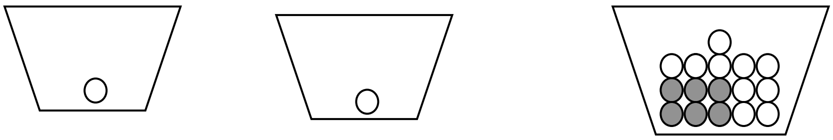


Fig. 17 : Une répartition des boules blanches et noires favorable au condamné à mort

Avec cette répartition, la probabilité de tirer une boule blanche vaut

$$P(B) = \frac{1}{3}\times 1 + \frac{1}{3}\times 1 + \frac{1}{3}\times\frac{10}{16} = \frac{42}{48} = \frac{21}{24} \approx 0{,}875.$$

Cette répartition accroît donc considérablement les chances de survie du condamné.

2.4 Deux problèmes avec deux urnes et des billes

L'urne I contient 4 billes rouges et 2 billes blanches ; l'urne II contient 3 billes rouges et 5 billes blanches. On fait deux expériences aléatoires différentes.

Première expérience :

On lance une pièce de monnaie. Si le résultat est « face », on tire une bille de l'urne I, si le résultat est « pile », on tire une bille de l'urne II. Quelle est la probabilité de tirer une bille rouge ?

Seconde expérience :

On tire une boule de chaque urne. Quelle est la probabilité de tirer

a) *deux billes blanches ?*
b) *deux billes rouges ?*
c) *une bille blanche et une bille rouge ?*

Ces deux expériences aléatoires permettent de s'exercer à la schématisation d'une expérience par un diagramme en arbre.

Dans la première expérience, les événements possibles sont : "choisir l'urne I et tirer une bille rouge", "choisir l'urne I et tirer une bille blanche", "choisir l'urne II et tirer une bille rouge", "choisir l'urne II et tirer une bille blanche".

Dans la seconde expérience, les événements possibles sont : "tirer une première bille rouge et tirer une seconde bille rouge", [tirer une première bille rouge et tirer une seconde bille blanche", "tirer une première bille blanche et tirer une seconde bille rouge", "tirer une première bille blanche et tirer une seconde bille blanche".

Voici maintenant les diagrammes en arbre correspondant respectivement à ces deux expériences (Fig.18).

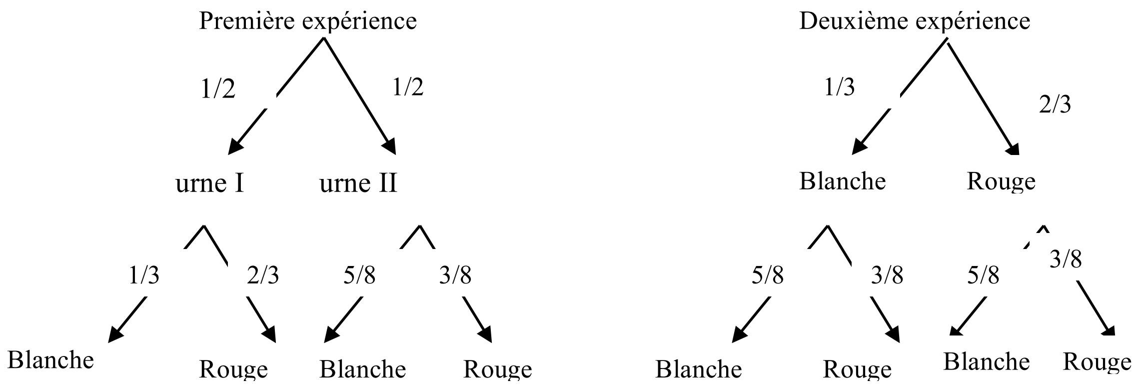


Fig.18 : Diagrammes en arbre probabilisé des deux expériences

Dans le cas de la première expérience, la probabilité de tirer une bille rouge est

$$\frac{1}{2}\times\frac{2}{3}+\frac{1}{2}\times\frac{3}{8}=\frac{25}{48}.$$

Dans le cas de la seconde expérience aléatoire, la probabilité de tirer deux billes blanches est

$$\frac{1}{3}\times\frac{5}{8}=\frac{5}{24},$$

la probabilité de tirer deux billes rouges est

$$\frac{2}{3}\times\frac{3}{8}=\frac{1}{4},$$

et la probabilité de tirer une bille blanche et une bille rouge est

$$\frac{1}{3}\times\frac{3}{8}+\frac{2}{3}\times\frac{5}{8}=\frac{13}{24}.$$

2.5 Répartition de la mise lors d'un jeu interrompu

Deux joueurs jouent à « pile ou face ». Le premier marque un point lorsqu'un lancer donne « face », le second lorsqu'un lancer donne « pile ». Le gagnant est le premier joueur qui arrive à six points. Il emporte la totalité de la mise fixée à 100 euros, chacun ayant misé 50 euros.

Un événement imprévu fait que le jeu est interrompu alors que le premier joueur mène 5 points à 3 pour le second. Comment faut-il répartir la somme mise en jeu ? Poursuivez le jeu à deux pour voir si la répartition que vous avez proposée semble correspondre à la réalité.

Puisque le jeu s'arrête sur la marque de 5 à 3, on pourrait partager les 100 euros proportionnellement aux points acquis : 5/8 de 100 euros = 62,5 euros pour le premier joueur et 3/8 de 100 euros = 37,5 euros pour le second.

On peut aussi se dire que pour gagner, le premier joueur n'a plus qu'un point à faire tandis que le second en a encore 3. On partagerait alors la mise en fonction des points qui manquent : (3/4) de 100 euros = 75 euros pour le premier joueur et (1/4) de 100 euros = 25 euros pour le second.

Dans ce dernier raisonnement, on considère que les quatre possibilités de finir le jeu ont toutes la même probabilité de se réaliser. Est-ce conforme à la réalité ? On pourrait le tester expérimentalement en observant un très grand nombre de parties, en ayant, par exemple, recours à une simulation grâce aux nombres aléatoires. On peut aussi calculer la probabilité de chacune des possibilités. Pour cela, commençons par représenter à la figure 19 le diagramme probabilisé du jeu, à partir du résultat obtenu au moment de son interruption. Dans ce diagramme, *p* représente la probabilité de gagner un point par le premier joueur et *q* représente la probabilité de gagner un point par le second joueur.

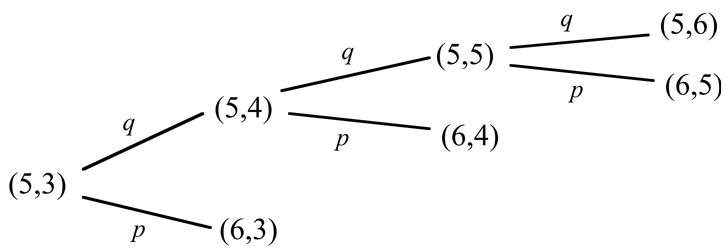


Fig. 19 : Diagramme en arbre de la partie du jeu à partir du résultat obtenu au moment de l'interruption du jeu

Ainsi, la probabilité du résultat (5, 6) est q^3, celle de (6,5) est q^2p, celle de (6,4) est qp, et celle de (6,3) est égale à p. Si la probabilité de gagner 1 point à chaque lancer de la pièce est la même pour chaque joueur (p = q = 0,5), ces quatre probabilités sont égales respectivement à $\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}$: les quatre issues ne sont donc pas équiprobables dans ce cas.

Il résulte aussi de ces calculs que la probabilité que le second joueur remporte la mise est égale à $\frac{1}{8}$ et que, par conséquent, la probabilité que le premier joueur remporte la mise est égale à $\frac{7}{8}$. Ainsi, le partage équitable serait de donner $\frac{7}{8}$ de 100 euros soit 87,5 euros au premier joueur et de donner $\frac{1}{8}$ de 100 euros soit 12,5 euros au second joueur.

Mais qui dit que les chances de gagner chaque partie sont les mêmes pour les deux joueurs? Vu la marque lorsque le jeu est interrompu, on peut croire que les probabilités de victoires respectives pour les deux joueurs sont de $p = \frac{5}{8}$ et $q = \frac{3}{8}$. Dans ce cas, la probabilité de victoire du second joueur vaut $\frac{3}{8} \times \frac{3}{8} \times \frac{3}{8} \approx 0{,}053$. Alors le partage équitable serait de donner au second joueur 5,5 euros et au premier joueur 94,7 euros.

2.6 Jeu de trois portes

Un jeu télévisé se présente comme suit : il y a trois portes fermées derrière lesquelles ont été disposées au hasard deux chèvres et une voiture. Le candidat désigne une porte de son choix. L'animateur du jeu n'ouvre pas cette porte, mais en ouvre une autre derrière laquelle se trouve une chèvre. Le candidat peut alors maintenir son premier choix ou de changer en désignant la porte restante. L'animateur ouvre la porte choisie et le candidat gagne l'objet qui se trouve derrière celle-ci.

À votre avis, que doit faire le candidat : maintenir son premier choix, changer de porte ou cela n'a pas d'importance ?

Selon certains, lorsque l'animateur ouvre l'une des portes derrière laquelle se trouve une de deux chèvres, il reste au candidat deux possibilités dont l'une est gagnante. Il a donc une chance sur deux de gagner la voiture.

Selon d'autres, lorsque l'animateur ouvre l'une des portes derrière laquelle se trouve une chèvre, il n'apporte aucune information utile au candidat; celui-ci a donc toujours une chance sur trois de gagner la voiture.

Qui a raison ? Pour mieux comprendre le fonctionnement du jeu, on peut jouer à deux ou faire une simulation avec un logiciel qui fournit des nombres aléatoires. Alors, on s'apercevra qu'à la deuxième étape, le joueur est obligé, soit de maintenir son choix, soit de le changer. Il s'agit donc de deux stratégies différentes.

Le tableau 7 reprend les résultats de 30 parties. Dans la première colonne, on donne la position (1, 2 ou 3) de la voiture ; dans la deuxième colonne, le choix du candidat et dans la troisième, la porte ouverte par l'animateur. Dans la quatrième colonne, on indique les succès (1) et échecs (0) du candidat lorsqu'il ne modifie pas son choix et dans la cinquième, les succès et échecs lorsqu'il le modifie. Prenons un exemple, à la première partie : la voiture se trouvait derrière la porte 3 et le candidat a choisi la porte 1. L'animateur a ouvert la porte 2 derrière laquelle se trouvait une chèvre. Si le candidat maintient son choix de départ, il perd. S'il le modifie et qu'il choisit la porte 3, il gagne.

À la dernière ligne, on a calculé la fréquence des succès lors des 30 parties. En gardant le premier choix, la fréquence des succès est de 0,4 et en le modifiant, elle est de 0,6. Mais on n'a joué que trente parties.

Voiture	Choix	Indication	Sans modifier	En modifiant
3	1	2	0	1
1	1	2	1	0
2	2	1	1	0
3	2	1	0	1
2	2	1	1	0
3	1	2	0	1
1	1	2	1	0
1	1	2	1	0
1	1	2	1	0
3	2	1	0	1
2	3	1	0	1
3	1	2	0	1
3	1	2	0	1
2	2	1	1	0
3	1	2	0	1
3	2	1	0	1
3	2	1	0	1
3	2	1	0	1
1	1	2	1	0
3	1	2	0	1
1	1	2	1	0
2	1	3	0	1
1	3	2	0	1
1	1	2	1	0
1	1	2	1	0
2	1	3	0	1
2	3	1	0	1
1	2	3	0	1
1	1	2	1	0
2	3	1	0	1
			0,400	0,600

Tabl.7

Les résultats de la simulation suggèrent qu'aucun des raisonnements proposés n'est correct. L'observation de ces résultats permet le raisonnement suivant. Chaque fois que le candidat désigne la voiture à son premier choix, il perd en le modifiant; chaque fois qu'il ne désigne pas la voiture à son premier choix, il gagne en le modifiant. Comme la probabilité de gagner au premier choix est de $\frac{1}{3}$, la probabilité de gagner en modifiant son choix est égale à $1 - \frac{1}{3} = \frac{2}{3}$.

Le déroulement du jeu peut être représenté par un diagramme en arbre (Fig.20).

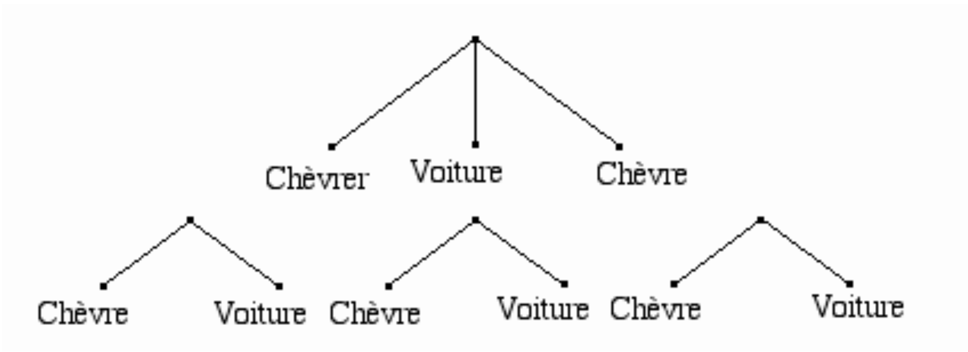


Fig. 20 : Diagramme des possibles associé au jeu de trois portes

Ce diagramme peut être probabilisé à deux manières différentes selon la stratégie choisie. Dans le cas de la stratégie « maintenir le premier choix », on obtient le diagramme de la figure 21.

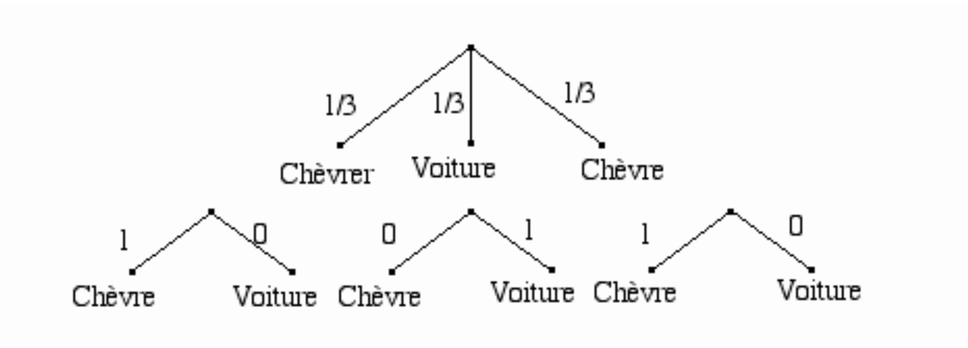


Fig. 21 : Diagramme probabilisé du jeu de trois portes associé à la 1e stratégie

Avec cette stratégie, la probabilité de gagner la voiture est égale à $\frac{1}{3}$.
Dans le cas de la stratégie « modifier le premier choix », on obtient le diagramme de la figure 22.

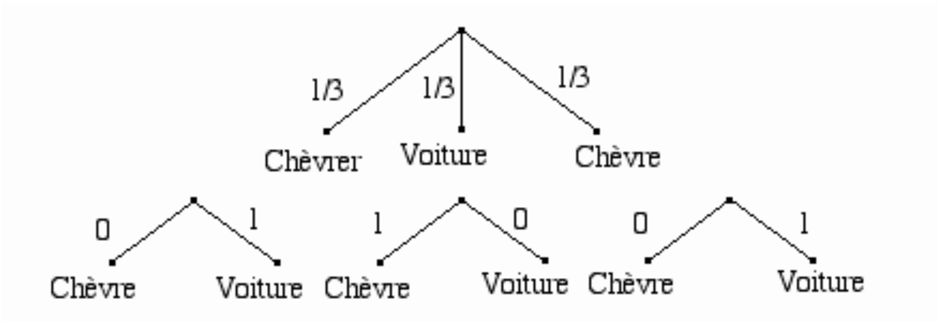


Fig.22 : Diagramme probabilisé du jeu de trrois portes associé à la 2e stratégie

Avec cette stratégie, la probabilité de gagner la voiture est égale à $\frac{2}{3}$.

En conclusion, à la deuxième étape, le joueur a intérêt à changer la porte qu'il avait choisie à la première étape. Mais en réalité, l'animateur, essaie d'influencer le joueur par des remarques du type « Est-ce votre dernier mot ? ». Si bien que la probabilité p de maintenir le premier choix peut prendre toutes les valeurs comprises entre 0 et 1. Et l'arbre probabilisé est celui de la figure 23.

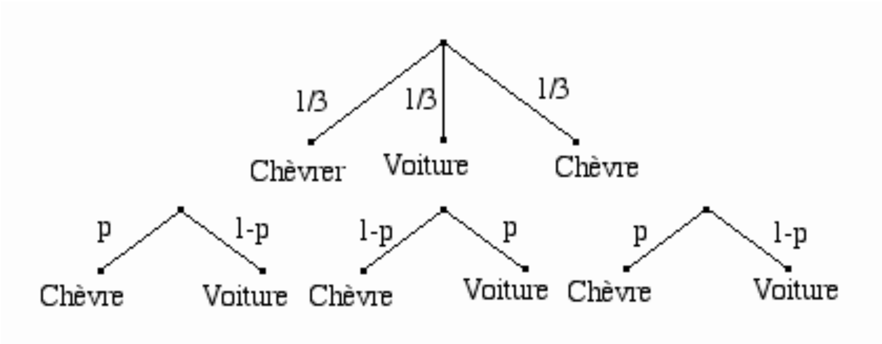


Fig.23 : Diagramme probabilisé du jeu de trois portes associé à une stratégie généralisé

De sorte que la probabilité de gagner la voiture est égale à

$$\frac{1}{3}(1-p)+\frac{1}{3}p+\frac{1}{3}(1-p)=\frac{2}{3}-\frac{1}{3}p.$$

Y a t-il une stratégie qui permet de gagner la voiture avec la probabilité de $\frac{1}{2}$? Pour y répondre il faut déterminer la valeur de p vérifiant l'égalité $\frac{2}{3}-\frac{1}{3}p=\frac{1}{2}$. Cela est vrai pour $p=\frac{1}{2}$. Cette stratégie consiste, à la deuxième étape, à choisir la porte en jouant à « pile » ou « face ». L'arbre probabilisé qui lui correspond est représenté à la figure 24.

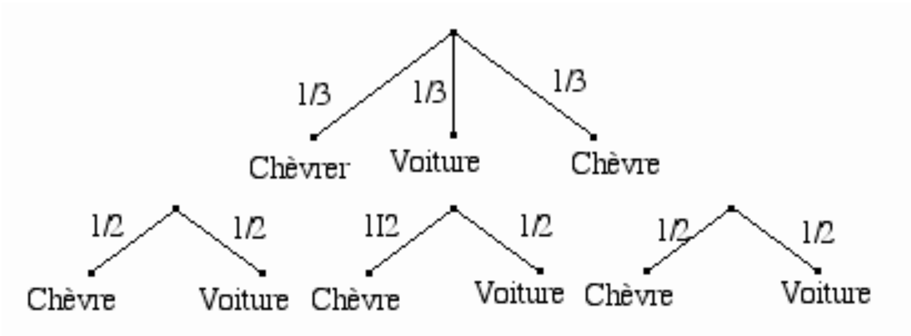


Fig. 24 : : Diagramme probabilisé du jeu de trois portes associé à la stratégie « pile » ou « face »

Vérifions à l'aide du diagramme que la probabilité de gagner la voiture est égale à

$$\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}=\frac{1}{2}.$$

Considérons le graphique de la fonction $f(p)=\frac{2}{3}-\frac{1}{3}p$ dans l'intervalle $[0,1]$ (Fig.25).

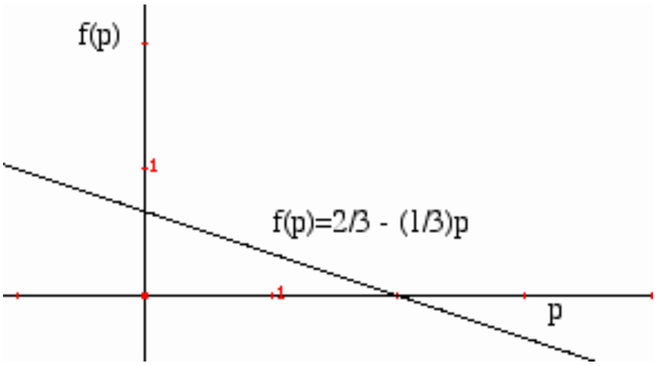


Fig. 25 : Graphique de la fonction $f(p) = \frac{2}{3} - \frac{1}{3}p$

Sur ce graphique, on observe que la probabilité de gagner est maximale et vaut 2/3 lorsque $p = 0$, c'est-à-dire lorsque le candidat change son premier choix.

Synthèse : Probabilités composées

Considérons des expériences aléatoires composées de plusieurs étapes. De telles expériences sont modélisées à l'aide des **diagrammes en arbres** dont la première fonction est de dénombrer tous les événements possibles au cours de la réalisation de l'expérience et la deuxième fonction est de donner des probabilités pour passer d'une étape à l'autre.

Un diagramme en arbre est caractérisé par des nœuds reliés par des chemins dont les règles d'utilisation s'appuient sur les règles d'utilisation des fréquences d'événements composés. Elles peuvent être énoncées de la manière suivante :

1. Le cheminement sur un arbre se fait uniquement de gauche à droite (si l'arbre est présenté horizontalement) ou de haut en bas (si l'arbre est présenté verticalement).
2. Chaque branche reliant deux nœuds successifs, est affectée de la probabilité de passer du premier nœud au second nœud.
3. La somme des probabilités affectées aux branches issues d'un même noeud est égale à 1.
4. La probabilité affectée à chaque chemin est le produit des probabilités affectées à chacune des branches qui le composent (**règle de multiplication**).
5. La probabilité d'un événement correspondant à plusieurs chemins de l'arbre est la somme des probabilités correspondant à chacun de ces chemins (**règle d'addition).**

Deux événements A et B rencontrés le long d'un chemin du diagramme en arbre sont des exemples d'événements réalisés simultanément. On peut alors considérer un nouvel événement appelé **conjonction des événements *A* et *B*** et noté **(*A* et *B*).** Citons comme exemple l'événement traité dans la section 2 .1 qui est le décès du couple pendant les dix ans de la durée d'un viager. Cet événement est la conjonction des deux événements : « le décès de l'homme avant l'âge de 85 ans » et « le décès de la femme avant l'âge de 80 ans ».

D'une manière générale, on dit que la conjonction (A et B) de deux événements A et B liés à une même expérience aléatoire est réalisée lorsque les deux événements sont réalisés simultanément.

Lors de la réalisation ou de la simulation d'une telle expérience répétée n fois, on calcule la fréquence de l'événement (A et B) de la manière suivante :

$$f_{A\ et\ B} = \frac{nombre\ de\ réalisations\ conjointes\ de\ A\ et\ B}{n}$$

où n est le nombre de répétitions de l'expérience.

Deux événements A et B réalisés sur plusieurs chemins du diagramme en arbre sont des exemples d'événements tels que soit l'un, soit l'autre, soit les deux sont réalisés. On peut alors considérer un nouvel événement appelé la **disjonction des événements *A* et *B*** et noté **(*A* ou *B*)**. Un exemple de tels événements a été traité dans la section 2.2, où on a calculé la probabilité P(3F et 1G ou 3G et 1F) de la disjonction des événements « une famille composée de 3 filles et d'un garçon « ou « une famille de 3 garçons et une fille »

D'une manière générale, on dit que la disjonction (A ou B) de deux événements liés à une même expérience aléatoire est réalisée lorsque au moins un des deux événements est réalisé.

Lors de la réalisation ou de la simulation d'une telle expérience répétée n fois, on calcule la fréquence de l'événement (A ou B) de la manière suivante :

$$f_{A\ ou\ B} = \frac{nombre\ de\ réalisations\ \ de\ A\ ou\ de\ B}{n}$$

où n est le nombre de répétitions de l'expérience.

La conjonction et la disjonction des événements sont des événements composés et leurs probabilités sont des **probabilités composées**. Pour les calculer, on utilise des diagrammes en arbres probabilisés.

Lorsque l'énoncé comporte des expressions comme *au moins, plus de, moins de, pas plus de…*, le calcul de la probabilité de l'événement contraire est parfois plus facile et plus rapide que celui de la probabilité demandée.

3 Probabilités conditionnelles

Dans ce chapitre, on se pose la question de savoir si une information concernant l'issue d'une expérience a une influence sur le calcul de la probabilité d'un événement associé à cette expérience. Ensuite, on y montre comment calculer la probabilité d'un événement sachant qu'un autre s'est produit.

3.1 Enquête de Test-Achat

Test-Achats a mené une enquête dans 86 garages de la province de Liège. Le tableau 9 donne les résultats de cette enquête. La qualité du service est-elle liée à la nature du garage (concessionnaire ou indépendant) ?

Service	Satisfaisant	Médiocre
Garagistes concessionnaires	18	6
Garagistes indépendants	34	28

Tab.8

Calculons la probabilité pour un client choisi au hasard, d'être satisfait de son garagiste : $\frac{18+34}{18+6+34+28} = \frac{52}{86} \approx 0{,}60$.

Calculons aussi la probabilité pour un client s'adressant à un concessionnaire, d'être satisfait de son garagiste: $\frac{18}{18+6} = \frac{18}{24} \approx 0{,}75$, et la probabilité pour un client s'adressant à un indépendant, d'être satisfait de son garagiste : $\frac{34}{34+28} = \frac{34}{62} \approx 0{,}55$.

Assimilons ces probabilités au degré de satisfaction d'un client vis-à-vis de son garagiste. Nous pouvons dire alors que d'après l'échantillon des clients testés, les garagistes concessionnaires donnent plus de satisfaction à leurs clients que les garagistes indépendants.

La probabilité d'être satisfait de son garagiste concessionnaire et la probabilité d'être satisfait de son garagiste indépendant sont des exemples de **probabilités conditionnelles**. Dans le premier cas, il s'agit de calculer la probabilité qu'un client soit satisfait sous condition que son garagiste est un concessionnaire ; dans le second cas, il s'agit de la probabilité qu'un client soit satisfait sous condition que son garagiste est un indépendant. Ajouter une telle condition modifie l'expérience aléatoire et donc ses résultats possibles. Dans le premier cas, il s'agit de tirer au hasard un client parmi ceux qui s'adressent à un garagiste concessionnaire ; dans le second cas, il s'agit de tirer au hasard un client parmi ceux qui s'adressent à un garagiste indépendant.

3.2 Pas bons les boulons

Une usine dispose de deux machines fabriquant des boulons. La première M_1 produit 40 % de la totalité des boulons et de plus 5 % des boulons fabriqués par cette machine sont défectueux. Tandis que la seconde ne produit que 1 % de boulons défectueux.

On tire, au hasard, un boulon produit par l'usine et l'on constate qu'il est défectueux, quelle est la probabilité qu'il ait été produit par la machine M_1 ?

Notons par

- M_1 l'événement « le boulon est fabriqué par la première machine »,
- M_2 l 'événement « le boulon est fabriqué par la deuxième machine' ,
- D l'événement « le boulon est défectueux »,
- non *D,* « le boulon est non défectueux ».

La réponse à cette question peut être exprimée aussi à l'aide des pourcentages :

- Le pourcentage de D produit par M_1 : 5% de 40% = 2%
- Le pourcentage de D produit par M_2 : 1% de 60% = 6‰

- Le pourcentage total de D produit par l'usine : 2% + 6‰ = 26‰

Et donc sur 26 boulons défectueux, M_1 en produit 20, ce qui représente $\frac{20}{26}$ soit environ 77% des boulons défectueux.

À partir des données, on peut construire un arbre probabilisé comme celui de la figure 26. Il correspond à l'expérience aléatoire composée consistant à tirer au hasard l'une des deux machines et, ensuite un boulon parmi les boulons fabriqués par celle-ci.

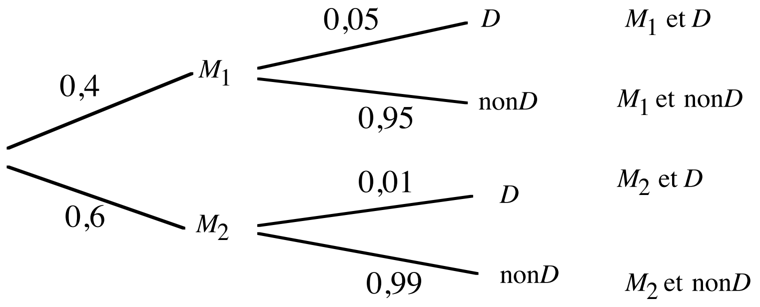


Fig. 26 : Diagramme probabilisé de la fabrication des boulons par deux machines

En appliquant les règles d'addition et de multiplication sur l'arbre de la figure 26, on obtient les probabilités suivantes :

$$P(M_1 \; et \; D) = 0{,}4{\times}0{,}05 = 0{,}02 = 2\%$$
$$P(D) = 0{,}4{\times}0{,}05 + 0{,}6{\times}0{,}01 = 0{,}026 = 2{,}6\%.$$

A partir des calculs effectués ci-dessus, on peut représenter le diagramme en arbre correspondant à l'expérience expérience aléatoire composée consistant à tirer au hasard un boulon parmi tous les boulons fabriqués par deux machines et ensuite choisir au hasard l'une des deux machines (Fig. 27).

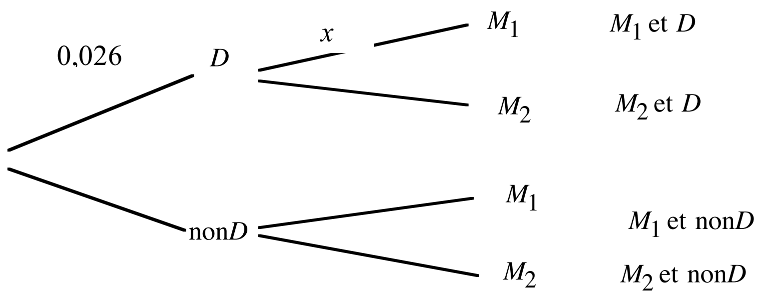


Fig. 27 Un autre diagramme en arbre de la fabrication des boulons par deux machines

À l'extrémité du premier chemins des arbres des figures 26 et 27, on rencontre le même événement composé : « un boulon défectueux **et** fabriqué par la machine M_1 », noté ($D \; et \; M_1$). Pour déterminer sa probabilité, on peut parcourir le premier chemin sur un arbre ou sur l'autre, mais le résultat doit être le même dans les deux cas.

En appliquant les règles d'addition et de multiplication sur l'arbre de la figure 27, on obtient

$$P(D \; et \; M_1) = x{\times}0{,}026.$$

Cette probabilité doit être égale à celle qu'on a obtenu sur l'arbre de la figure 26 ce qui nous donne

$$P(D\ et\ M_1) = x \times 0{,}026 = 0{,}02$$

d'où

$$x = \frac{0{,}02}{0{,}026} \approx 0{,}77 = 77\%.$$

A partir de tous ces calculs on peut conclure que la probabilité de tirer un boulon défectueux parmi les boulons produits par la première machine est plus grande que la probabilité de tirer un boulon défectueux parmi tous les boulons produits par l'usine :

$$P(D/M_1) = 0{,}026 \text{ et } P(D) = 0{,}02.$$

La différence entre ces probabilités tient au fait que dans $P(D/M_1)$, le référentiel n'est pas le même que dans $P(D)$. Dans le premier cas, on considère les boulons produits seulement par la machine M_1 ; dans le second cas, on considère tous les boulons produits par l'usine.

On peut aussi conclure que la probabilité qu'un boulon défectueux soit produit par la première machine est très grande et qu'elle n'est pas la même que celle d'un boulon défectueux produit par l'usine :

$$P(M_1/D) = 0{,}77 \text{ et } P(D) = 0{,}02$$

Dans le cas de $P(M_1/D)$, il y a une information supplémentaire qui modifie le référentiel de $P(M_1)$. En effet, le référentiel de $P(M_1)$ est composé de tous les boulons produits par l'usine; tandis que le référentiel de $P(M_1/D)$ est composé seulement les boulons défectueux.

La probabilité qu'un boulon soit défectueux s'il est produit par la première machine est un exemple de **probabilité conditionnelle** ; on parle de **probabilité de $\boldsymbol{D}$ si $\boldsymbol{M_1}$** que l'on note $\boldsymbol{P(D/M_1)}$**.**

De même, la probabilité que la machine M_1 **produise un boulon sachant qu'il est défectueux** est un autre exemple de probabilité conditionnelle **;** on parle de **la probabilité de $\boldsymbol{M_1}$ si $\boldsymbol{D}$** que l'on note $\boldsymbol{P(M_1/D)}$**.**

A l'aide des probabilités conditionnelles, on peut reformuler la règle de multiplication associées au diagramme en arbre de la figure 26 sous la forme

$$P(M_1\ et\ D) = P(M_1) \times P(D/M_1), \qquad (1)$$

et la règle d'addition sous la forme

$$P(D) = P(M_1) \times P(D\ /\ M_1) + P(M_2) \times P(D/M_2).$$

Pareillement pour le diagramme de la figure 27, auquel on peut associer la règle de multiplication reformulée :

$$P(M_1\ et\ D) = P(\ D\ et\ M_1) = P(D) \times P(M_1/D) \qquad (2)$$

La structure identique des formules (1) et (2) nous donne la définition d'une probabilité conditionnelle :

$$P(A/B) = \frac{P(A\ et\ B)}{P(B)}$$

à condition que $P(B) \neq 0$.

3.3 Boules blanches et noires

Une urne contient deux boules blanches et deux boules noires. On les mélange et l'on tire l'une après l'autre, deux boules sans les remettre.

a) Quelle est la probabilité que la deuxième boule soit blanche si la première boule est blanche ?

b) Quelle est la probabilité que la première boule soit blanche si la seconde boule est blanche ?

a) La réponse à cette question ne pose aucun problème: étant donné qu'on a déjà tiré une boule blanche, il reste dans l'urne une boule blanche et deux boules noires. Ainsi il y a une chance sur trois de prendre la boule blanche au deuxième tirage.

b) En général, on pense que le résultat du deuxième tirage ne peut plus influencer la probabilité d'avoir la boule blanche au premier tirage. C'est pour cette raison que beaucoup maintiennent l'avis selon lequel on garde une chance sur deux de tirer une boule blanche au premier tirage même si, au deuxième tirage, on a aussi tiré une boule blanche.

Pour nous convaincre que ce raisonnement est faux, envisageons momentanément la situation suivante: dans cette urne, on tire l'une après l'autre, trois boules sans les remplacer. On ne regarde pas la couleur de la première boule, mais on le fait pour les deux autres boules et l'on constate qu'elles sont blanches. Dans cette situation va t-on maintenir l'avis selon lequel il y a une chance sur deux de tirer une boule blanche au premier tirage ? Bien sûr que non, on dira même qu'il n'y a aucune chance d'avoir une boule blanche au premier tirage. Donc le résultat des deux autres tirages a modifié la probabilité d'avoir la boule blanche au premier tirage.

Revenons donc à la situation décrite en b) et envisageons une simulation de tirages de deux boules sans remise. Notons par « B_1 » l'événement « la première boule tirée est blanche », par « B_2 » l'événement « la deuxième boule tirée est blanche », par « N_1 » l'événement « la première boule tirée est noire » et par « N_2 » l'événement « la deuxième boule tirée est noire ». On constate expérimentalement que les séquences B_1 B_2 sont plus rares que les séquences N_1B_2. Cela permet de conjecturer que, dans le cas où le deuxième tirage aurait donné une boule blanche, la probabilité d'avoir une boule blanche au premier tirage est plus petite que celle d'avoir une boule noire.

À partir d'un diagramme en arbre présenté à la figure 28, calculons cette probabilité en utilisant la définition de probabilité conditionnelle.

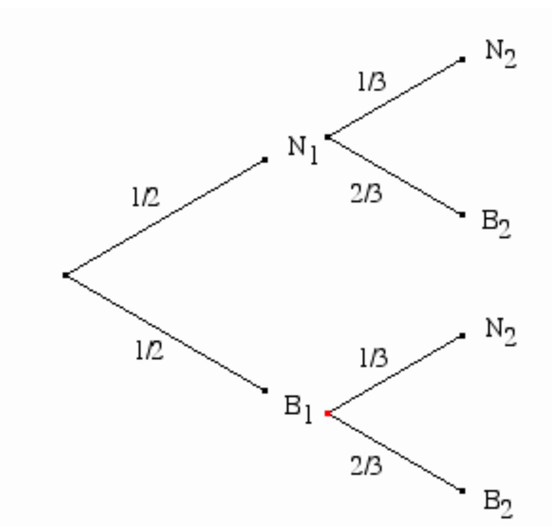


Fig. 28 : Diagramme probabilisé en arbre du tirage de deux boules

Cela donne

$$P(B_1/B_2) = \frac{P(B_1\ et\ B_2)}{P(B_2)} = \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{3}+\frac{1}{2}\times\frac{2}{3}} = \frac{1}{3}.$$

PP

Lors de l'expérimentation ou de la simulation, on voit aussi que pour estimer une probabilité conditionnelle à l'aide des fréquences, il faut considérer une autre expérience aléatoire qui consiste à ne retenir que les issues de l'expérimentation où le deuxième tirage a donné une boule blanche. En termes de probabilités, on change le référentiel ce qui revient à ne garder sur l'arbre de la figure 28 que les chemins qui mènent à l'événement « la deuxième boule tirée est blanche ». On s'intéresse alors à un arbre partiel (Fig.29) ce qui nous oblige à mettre une autre répartition sur les chemins restants. Puisque la séquence N_1B_2 avait deux fois plus de chances de se produire que la séquence B_1N_2, cette répartition doit être 2/3 pour N_1B_2 et 1/3 pour B_1N_2. La probabilité d'avoir une boule blanche au premier tirage sachant, qu'au deuxième tirage, on a eu aussi une boule blanche, est donc égale à 1/3.

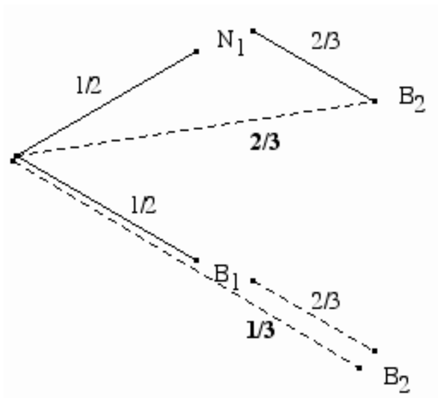


Fig. 29 : Diagramme probabilisé en arbre partiel du tirage de deux boules lorsque la deuxième boule tirée est blanche

Ce problème montre que la probabilité conditionnelle n'a rien à voir, ni avec un lien de causalité (cause à effet), ni avec un lien chronologique (passé au présent).

3.4 Efficacité d'un test (inspiré par [J. Lubczanski, 1994])

Depuis quelque temps une nouvelle maladie se répand et l'on estime qu'aujourd'hui 1% de la population est atteint. Un laboratoire a mis au point un test de dépistage. Mais aucun test n'est efficace à 100 %, il y a toujours un pourcentage d'erreur.

Pour tester l'efficacité du test, on l'a expérimenté sur des individus dont l'état est connu :

- sur un échantillon d'individus malades, le test a été positif pour 99 % d'entre eux;

- sur un échantillon d'individus non-malades, le test a été négatif pour 98 % d'entre eux.

Quelle est l'efficacité du test pour quelqu'un dont on ne connaît pas l'état de santé ?

Notons par « *M* » l'événement « un individu est malade », par « *S* » l'événement « un individu est sain », par « *P* » l'événement « le test est positif » et par « *N* » l'événement « le test est négatif ». Convenons de considérer qu'un test est efficace si $P(M/N)$ est inférieur à 0,001.

Les données du problème peuvent être représentées sous forme d'un diagramme en arbre probabilisé (Fig.30).

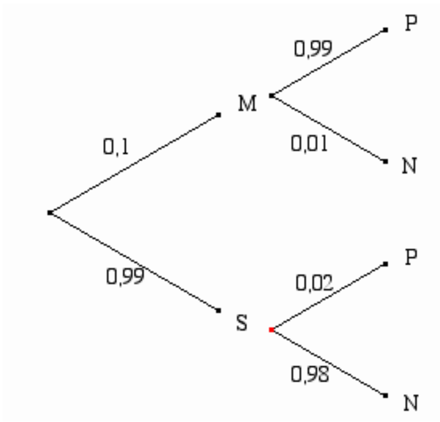


Fig. 30 : Diagramme probabilisé en arbre d'efficacité d'un test

Calculons $P(M/N)$:

$$P(M/N) = \frac{P(M\ et\ N)}{P(N)} = \frac{0{,}01\times 0{,}01}{0{,}01\times 0{,}01+0{,}00\times 0{,}98} = \frac{0{,}0001}{0{,}970298} = 0{,}0001030611.$$

Cette probabilité est quasi nulle. On considère que le test est fiable : parmi 10 000 personnes ayant eu le test négatif, il y en aura à peu près une qui contractera la maladie.

3.5 Deux perroquets (inspiré par [Gardner (1980)])

Madame Piat possède deux perroquets, un blanc et un noir.

a) Un jour, un visiteur lui demande: « L'un deux est-il un mâle ? ». « Oui » répond-t-elle. Quelle est la probabilité que les deux perroquets soient des mâles ?

b) Un autre jour, un autre visiteur lui demande: « Le blanc est-il mâle ? ». « Oui », répond-t-elle. Quelle est pour vous la probabilité que les deux oiseaux soient des mâles ?

Désignons par « *MB* » et « *MN* » le fait qu'un perroquet mâle soit respectivement blanc et noir, par « *FB* »et « *FN* » le fait qu'un perroquet femelle soit respectivement blanc et noir.

Après avoir convenu qu'à la naissance, il y a autant de perroquets mâles que de femelles, construisons l'arbre des probabilités (Fig.31).

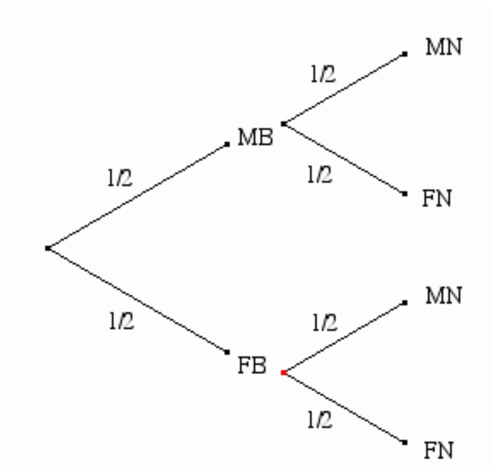


Fig. 31 : Diagramme en arbre probabilisé du problème de deux perroquets

a) Appelons « A » l'événement « Mme Piat possède au moins un mâle ». À partir de la figure 32 et en utilisant la définition de probabilité conditionnelle, on a

$$P(2\ m\hat{a}les/A) = \frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} = \frac{1}{3}.$$

Cela revient à ne pas considérer le dernier chemin de l'arbre de la figure 31 et à ne garder que les 3 premiers chemins. Cette nouvelle expérience décrite par l'arbre partiel de la figure 32, nous oblige à mettre une autre répartition sur les chemins restants, celle-ci ne peut être que 1/3, 1/3 et 1/3 puisque ces chemins restent équiprobables. La probabilité demandée est donc égale à 1/3.

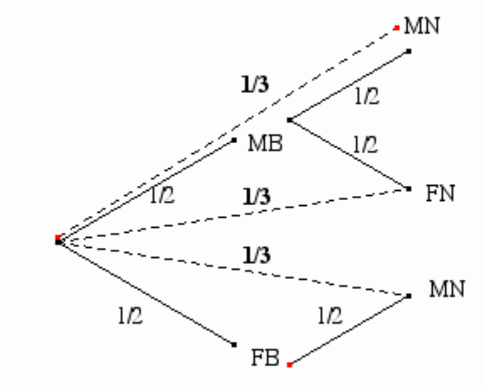


Fig. 32 :Diagramme en arbre probabilisé partiel lorsqu'au moins un perroquet est mâle

b) Appelons « B » l'événement « Mme Piat possède un perroquet mâle qui est blanc ».
À partir de la figure 32 et en utilisant la définition de probabilité conditionnelle, on a

$$P(2\ m\hat{a}les/B) = \frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} = \frac{1}{2}.$$

Cela revient à ne garder que les deux premiers chemins de l'arbre de la figure 32. Cette nouvelle expérience décrite par l'arbre partiel de la figure 33 nous oblige à mettre une autre répartition sur les chemins restants, celle-ci ne peut être que 1/2, 1/2, puisque ces chemins restent équiprobables. La probabilité recherchée est donc égale à 1/2.

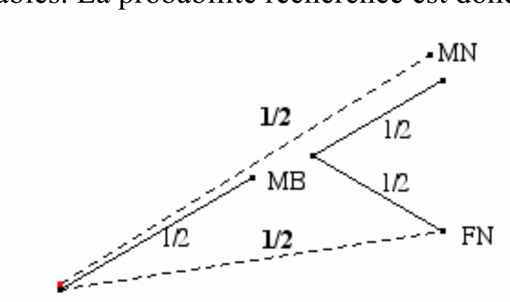


Fig.33 : Le diagrameen arbre probabilisée partiel lorsque le perroquet blan est mâle

De nouveau, on peut voir ici que les informations supplémentaires modifient le référentiel initial et donc le calcul des probabilités.

3.6 Trois jetons

Dans un sac, il y a trois jetons, l'un avec deux faces noires, l'autre avec deux faces blanches et le dernier avec une face blanche et une face noire. On tire un jeton au hasard, on regarde une de ses faces: elle est blanche. Quelle est la probabilité que l'autre face soit aussi blanche?

L'expérience aléatoire consiste à choisir au hasard d'abord un jeton et ensuite choisir au hasard la face du jeton.

Notons par « BB », « BN » et « NN » les jetons ayant respectivement 0,1 et 2 faces noires; par « B_1 » et « B_2 » les faces du jeton « BB ». Notons encore par « N_1 » et « N_2 » les faces du jeton « NN », et par « N » et « B » les faces du jeton « BN ».
Ce qui nous mène à l'arbre de la figure 34.

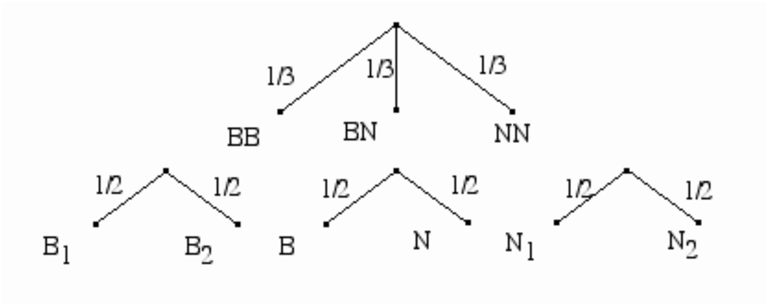


Fig. 34 : Diagramme en arbre probabilisé du problème de trois jetons

À partir de la figure 34 et en utilisant la définition de la probabilité conditionnelle on obtient

$$P(deuxi\grave{e}me\ face\ blanche/une\ face\ blanche) = \frac{\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}}{\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}} = \frac{2}{3}.$$

Autrement dit, cette expérience décrite par l'arbre partiel de la figure 35 nous oblige à mettre une autre répartition sur les chemins restants, celle-ci ne peut être que 1/3, 1/3 et 1/3 puisque ces chemins restent équiprobables.

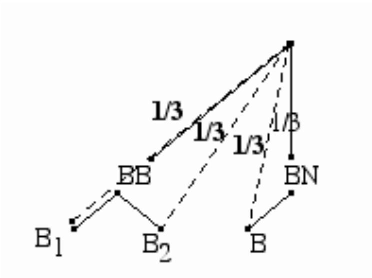


Fig. 35 : Le diagramme en arbre partiel lorsqu'une face est blanche

La probabilité demandée est donc égale à 2/3 puisque deux chemins réalisent l'événement « deux faces blanches ».

Synthèse : Probabilités conditionnelles

Soient deux événements A et B associés à une même expérience aléatoire. Nous définissons la **probabilité conditionnelle** de l'événement B sous condition que l'événement A s'est produit, de la manière suivante:

$P(B/A) = \frac{P(A\ et\ B)}{P(A)}$.

Cela signifie qu'on ne retient que les événements favorables à A et parmi ceux-là on comptabilise ceux qui sont aussi favorables à B.

Lors d'une simulation d'un événement conditionnel, on calcule la fréquence de cet événement de la manière suivante :

$f_{B/A} = \frac{nombre\ de\ résultats\ qui\ réalisent\ A\ et\ B}{nombre\ de\ résultats\ qui\ réalisent\ A}$.

La définition de la probabilité conditionnelle est cohérente avec la règle du produit le long d'un chemin du diagramme en arbre. En effet, les branches d'un tel diagramme correspondent à des probabilités conditionnelles (Fig.35) et la règle du produit le long d'un chemin s'écrit

$P(A)\times P(B/A) = P(B\ et\ A)$.

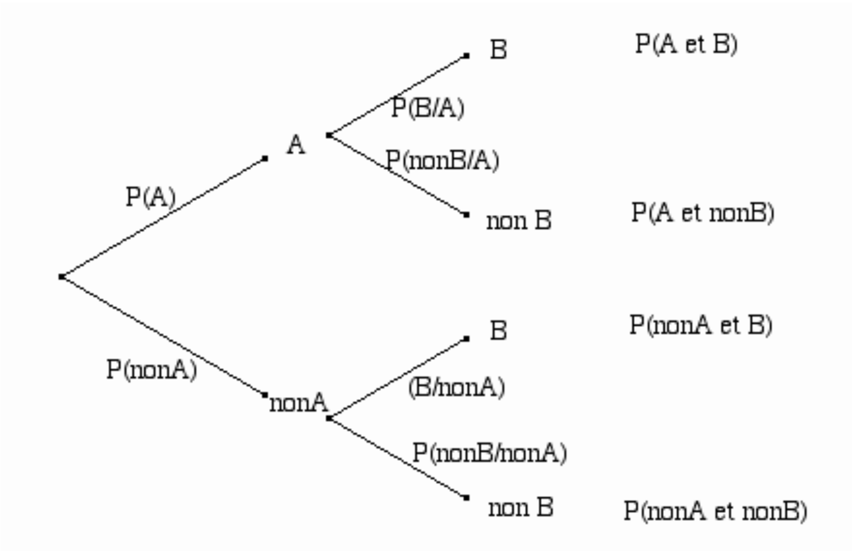


Fig. 36 ; Diagramme en arbrte avec les probabilitézs conditionnelles

.

4 Indépendance en probabilité

Dans ce chapitre, nous allons étudier deux événements indépendants en probabilité, c'est-à-dire des événements tels que la réalisation de l'un d'eux n'affecte pas la probabilité de l'autre événement.

4.1 Tabagisme (d'après [Engel 1975])

Les employés de deux usines sont classés selon deux critères : leur sexe et leur comportement vis-à-vis du tabac.

Les résultats obtenus dans deux usines sont présentés dans les tableaux 9a et 9b.

Usine 1	Femmes	Hommes	
Fumeurs	200	600	800
Non fumeurs	100	300	400
	300	900	1200

Tabl.9a

Usine 2	Femmes	Hommes	
fumeurs	200	800	1000
Non fumeurs	300	200	500
	500	1000	1500

Tabl. 9b

Pour chacune des deux usines, on choisit une personne au hasard.

a) *Calculer la probabilité de choisir une femme. Cette probabilité reste-t-elle la même si on sait que la personne choisie est fumeur?*
b) *Calculer la probabilité de choisir une personne fumeur. Cette probabilité reste-t-elle la même si on sait que la personne choisie est une femme?*

Notons par « F » l'événement « être femme » et par « Fu » l'événement « être fumeur ». Pour chaque usine, on peut construire deux arbres de probabilités. Les figures 37a et 37b présentent les arbres relatifs à la première usine.

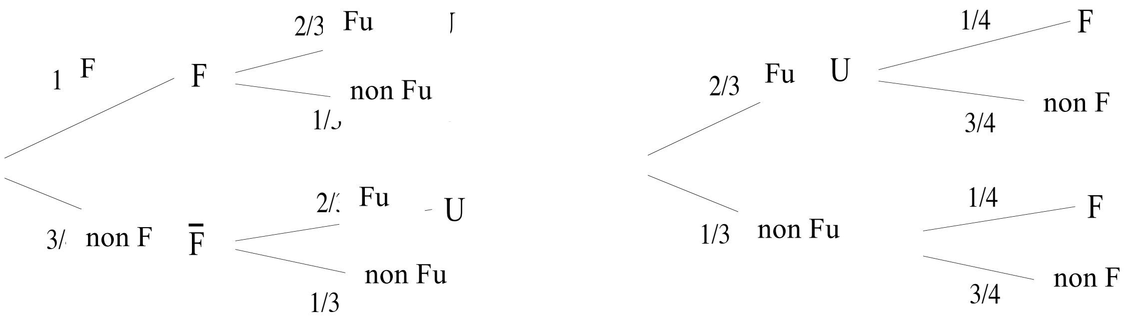


Fig. 37 : Diagramme en arbre probabilisé des fumeurs et des non fumeurs de la 1e usine

Calculons les probabilités demandées relatives à la première usine :

$$P(F) = 1/4 = P(F/ Fu),$$
$$P(Fu) = 2/3 = P(Fu / F).$$

Dans cette usine, les probabilités des événements « être femme » et « être femme sachant qu'on est fumeur » sont égales. Il en est de même pour les événements « être fumeur » et « être fumeur sachant qu'on est femme ». Que l'on considère d'abord les événements « être femme » ou non avant les événements « être fumeur » ou non, ou qu'on les considère dans l'ordre inverse, c'est la même chose. Dans ce cas, on dit que les événements sont **indépendants**.

Les figures 38a) et 38b) présentent les arbres relatifs à la seconde usine.

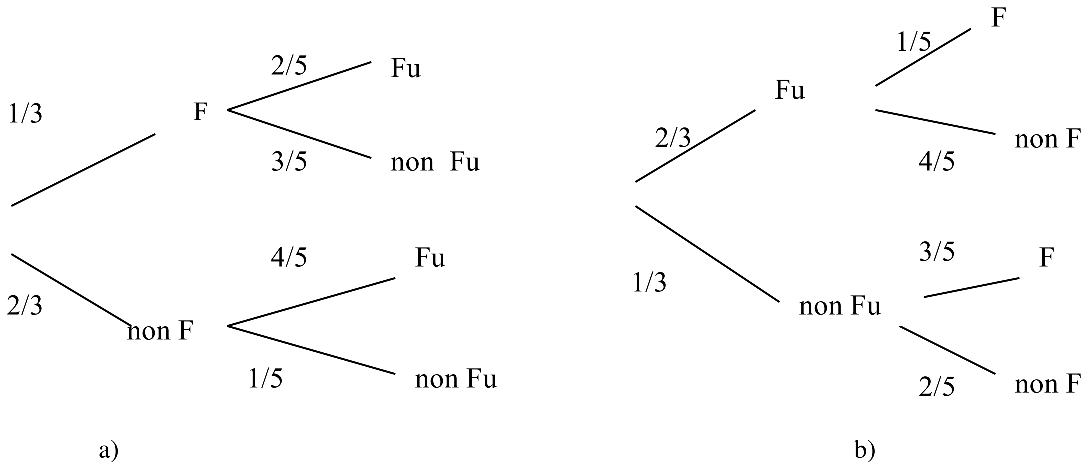


Fig. 38 : Diagramme en arbre probabilisé des fumeur et des non fumeurs de la 2e usine

Calculons les probabilités demandées relatives à la deuxième usine

$$P\ (F) = 1/3 \neq P\ (F/\ Fu) = 1/5,$$
$$P\ (Fu) = 2/3 \neq P\ (Fu/\ F) = 2/5.$$

On constate que dans cette usine, les événements « être femme » et « être femme sachant qu'on est fumeur » n'ont pas la même probabilité. Il en est de même pour les événements « être fumeur » et « être fumeur sachant qu'on est femme ». Que l'on considère les événements « être femme » ou non avant les événements « être fumeur » ou non, ou que qu'on les considère dans l'ordre inverse, ce n'est pas la même chose. Dans ce cas, on dit que les événements sont **dépendants**.

Remarquons, au passant, que le tableau de la première usine (Tabl.10a) est un tableau de proportionnalité et que, en même temps, les événements F et Fu sont indépendants. Par contre, le tableau de la deuxième usine (Tabl.10b) n'est pas un tableau de proportionnalité. et que , en même temps, les événements F et Fu sont dépendants.

4.2 Avec ou sans remise

Une urne contient 10 boules de forme et d'aspect identiques sauf en ce qui concerne la couleur, 6 d'entre elles sont rouges et les autres blanches.

Appelons R l'événement « la boule tirée est rouge », A l'événement « la première boule tirée est rouge » et B l'événement « la seconde boule tirée n'est pas rouge ».

a) Lorsqu'on tire deux boules au hasard, sans remettre la première avant de tirer la seconde, les événements A et B sont-ils indépendants ?
b) Lorsqu'on tire deux boules au hasard en remettant la première avant de tirer la seconde, les événements A et B sont-ils indépendants ?

a) Dans le cas d'un tirage sans remise, la proportion des boules rouges et blanches avant le deuxième tirage varie selon le résultat du premier tirage. On peut raisonnablement supposer que la probabilité de l'événement *B* dépend de la réalisation ou non de l'événement *A*. Pour le confirmer, modélisons cette expérience à l'aide d'un arbre (Fig.39) :

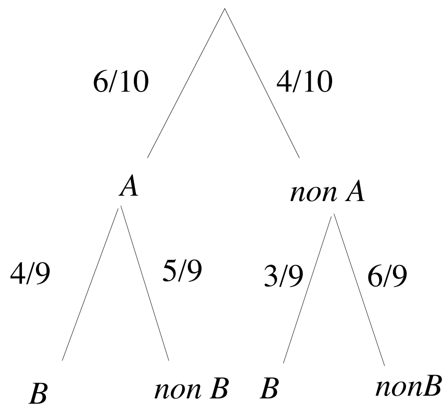


Fig. 39 :Diagramme en arbre probabilisé « avec remise »

On a

$$P(B|A)=\frac{4}{9} \text{ et } P(B)=\frac{6}{10}\times\frac{4}{9}+\frac{4}{10}\times\frac{3}{9}=\frac{4}{10}.$$

Ainsi $P(B|A)\neq P(B)$ et donc les événements A et B sont dépendants.

b) Dans le cas d'un tirage avec remise, la proportion des boules rouges et blanches reste la même avant chaque tirage suivant. On peut donc raisonnablement supposer que les événements A et B sont indépendants. Les calculs à partir de l'arbre associé le confirment (Fig.40).

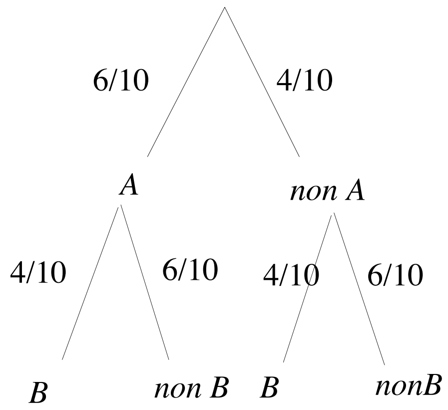


Fig. 40 : Diagramme en artbre probabilisé « sans remise »

On a

$$P(B|A)=\frac{4}{10} \quad \text{et} \quad P(B)=\frac{6}{10}\times\frac{4}{10}+\frac{4}{10}\times\frac{4}{10}=\frac{4}{10}.$$

Ainsi $P(B|A)=P(B)$ et donc les événements A et B sont indépendants.

On remarque ici que, dans ce cas, les arbres partiels du deuxième étage sont identiques; c'est ainsi que se traduit le tirage avec remise : il est à l'origine de l'indépendance des événements A et B.

4.3 Famille de trois enfants (tiré de [Lipschutz S (1973]),

Parmi les familles de trois enfants, les événements «avoir des enfants des deux sexes» (noté A) et «avoir au plus une fille»(noté B) sont-ils indépendants ?

Au total il y a huit types de familles possibles :

FFF, FFG, FGF, FGG, GFF, GFG, GGF, GGG.

L'événement A est constitué des six types de familles suivantes :

FFG, FGF, FGG, GFF, GGF, GFG.

L'événement B est constitué des quatre types de familles suivantes :

FGG, GFG, GGF, GGG.

L'événement A et B est constitué des trois types de familles suivantes :

FGG, GFG, GGF.

a) Si les naissances des garçons et des filles sont considérées comme équiprobables, tous les types de familles sont également équiprobables, chacun ayant la probabilité de 1/8 (Section 2.2). D'où

$$P(A/B) = \frac{P(A\ et\ B)}{P(B)} = \frac{\frac{3}{8}}{\frac{4}{8}} = \frac{3}{4} = \frac{6}{8} = P(A).$$

Il en résulte que les événements A et B sont indépendants.

b) Supposons maintenant que la probabilité de naissance d'un garçon est 0,5135 et celle d'une fille est 0,4865 (Section 2.1). Alors, d'une part

$$P(A) = 3\times 0{,}4865^2 \times 0{,}5135 + 3\times 0{,}5135^2 \times {,}4865 \approx 0{,}7495$$

et, d'autre part

$$P(A/B) = \frac{3\times 0{,}5135^2 \times {,}4865}{3\times 0{,}5135^2 \times {,}4865 + 0{,}5135^3} \approx 0{,}7397.$$

Les deux probabilités $P(A)$ et $P(A/B)$ sont différentes donc les événements A et B sont, cette fois-ci, dépendants. Cela signifie qu'un petit changement dans la répartition des probabilités des événements donnés transforme leur indépendance en dépendance. Cela montre aussi que l'indépendance de deux évènements n'est pas leur propriété intrinsèque, qu'elle dépend du modèle probabiliste choisi.

Synthèse : Indépendance

Soient deux événements A et B liés à une même expérience aléatoire. L'événement A **est indépendant de l'événement B** si la réalisation de B ne modifie pas la valeur de la probabilité de la réalisation de A c'est-à-dire si

$$P(A/B) = P(A)$$

1. Si A est indépendant de B, alors B est indépendant de A

En effet, par hypothèse, on a

$$P(A/B) = P(A)$$

Et par définition, on a aussi

$$P(A/B) = \frac{P(A\ et\ B)}{P(B)}.$$

De là, on obtient

$$P(A) = \frac{P(A\ et\ B)}{P(B)},$$

ce qui est équivalent à

$$P(B) = \frac{P(A\ et\ B)}{P(A)} = (B/A).$$

La symétrie de la relation d'indépendance d'un événement par rapport à un autre permet dorénavant de parler de deux **événements indépendants en probabilité**.

2. Si deux événements A et B sont indépendants, alors $P(A\ et\ B) = P(A)P(B)$

Cette relation se déduit immédiatement de la définition de la probabilité conditionnelle. Elle peut être utilisée comme un critère pour décider de l'indépendance de deux événements.

3. Si les événements A et B sont indépendants, alors les événements *nonA* et B le sont aussi.

En effet, comme les événements A et B, *non A* et B correspondent aux deux branches issues d'un même nœud dans un arbre probabilisé, on a

$$P(B\ et\ A) + P(B\ et\ nonA) = P(B).$$

Ainsi,

$$P(nonA/B) = \frac{P(B\ et\ nonA)}{P(B)} = \frac{P(B)-P(B\ et\ A)}{P(B)} = 1 - P(A/B) = 1 - P(A) = P(nonA).$$

4. Des événements incompatibles (notamment des événements contraires) sont dépendants

En effet, la réalisation de l'un implique la non réalisation de l'autre, sauf dans le cas où l'un d'eux est l'événement impossible.

Bibliographie

Engel A. (1975), *L'enseignement des probabilités et de la statistique,* vol 1, ed CEDIC, Paris

Engel A. (1975), *L'enseignement des probabilités et de la statistique,* vol 2, ed CEDIC, Paris

Fischbein & Schnarch, (1997),The evolution with age of probabilistic Intuitivly based misconceptions in *Journal for Research in Mathematics Education,* Vol 28, No. 1 pp. 96-105

Gardner M.,(1980) *La magie des paradoxes*, Bibliothèque pour la science - diffusion Belin

Lipschutz S., (1973), *Probabilités,* Serie Schaum, McGraw-Hill

Lubczanski J. (1994), *Le trésor de tonton Lulu*, Vol. 2, eds Archimède Eds., Diffusion Tangente, (A.P.M.E.P)