LE BIATHLON, L'AIR DE JAZZ, LES OASIS, LA BELLE AU BOIS DORMANT, LEURS UNIVERS ET LEURS ENTROPIES

Philippe Gay

Résumé. Quatre paradoxes probabilistes, dont trois inédits, sont soumis à la sagacité du lecteur. Les différentes démonstrations sont communes à l'ensemble de ces énigmes et utilisent la notion de base d'univers, des diagrammes en treillis et arborescents et des notations matricielles. La méthode se prête bien à des simulations diverses, accessibles au monde de l'enseignement, et illustre la notion de glissement bayésien. Il apparaît qu'une probabilité peut prendre différentes valeurs si l'on considère des espaces de probabilités différents. Une généralisation de ces problèmes à l'aide de mélange d'urnes est proposée.

Mots-clés. Probabilité conditionnelle, Bayes, événement, univers, espace de probabilités, simulation, paradoxe, graphe en treillis, graphe arborescent, Belle au bois dormant.

Abstract. Four probability paradoxes, among which three are unpublished works, are subjected to the reader's sagacity. Several demonstrations, common to all these drills, use the basic notion of universe, wire-mesh and tree diagrams and matrix notations. These methods are well suited to diverse simulations, easy-to-use for education, and can illustrate the Bayesian-sliding notion. It appears that a probability can take different values if we consider different universes. A generalization of these problems is proposed by combining urns.

Keywords. Conditional probability, Bayes, event, universe (or sample space), probability space, simulation, paradox, Sleeping Beauty.

1 Les Paradoxes : un défi

Peut-on voir un lien entre les différentes histoires décrites ci-dessous, dont les trois premières sont inédites et la quatrième connue depuis plusieurs années ? Oui, si l'on définit correctement leurs bases mathématiques. On pourra même étendre la méthode à quantité de problèmes utilisant les probabilités conditionnelles. Une formule générale sera proposée.

Le biathlon est une épreuve olympique regroupant ski de fond et tirs à la carabine. Martin est justement sur le pas de tir et cela ne se passe pas bien. Son entraineur téléphone à son père Marcel et lui confie : *« Allez plutôt sur la boucle réservée à ceux qui ont manqué une ou plusieurs cibles. Martin devra y faire un ou deux tours de pénalités. »* On suppose que Martin a une chance sur deux de faire un tour de pénalité et de même pour deux tours de pénalité (suivant le résultat sur sa dernière cible). Envisageons deux cas eux aussi équiprobables. Si Marcel arrive tôt sur la boucle, il assistera au premier tour de piste. S'il arrive plus tard (et si Martin est suffisamment pénalisé), Marcel ne verra au mieux que le second tour. Marcel arrive sur cette piste et voit son fils passer. *De son point de vue, quelle est la probabilité qu'il soit en train de voir Martin sur son premier (ou son second) tour de pénalité ?*

Un air de Jazz (*Take Five*) : Fabienne supporte difficilement la musique de son père Pierre qui dispose de deux versions de *Take Five* : la version courte de 5 minutes et celle plus longue de 10 minutes. Pour Fabienne il n'est possible ni de distinguer le début de la fin, ni de distinguer la version courte de la version longue. Juste elle peut dire : *« C'est* Take Five *! »* ou pas. Sa culture du jazz s'arrête à cela. On suppose que Pierre choisit la version courte une fois sur deux (et donc de même pour la version longue). Fabienne arrive à proximité de l'appartement et elle ne sait pas si le morceau de musique est commencé depuis longtemps. On suppose que l'arrivée de Fabienne en première moitié ou seconde moitié de la version longue sont équiprobables. *Quelle est la probabilité que Fabienne entende la version courte ?*

Les Oasis : Jennifer et Maurice profitent de leurs vacances au Kesaltan pour visiter les oasis. Jennifer hésite entre la première formule où l'on voit 2 beaux oasis et celle plus longue où l'on en voit 4. *« Je donnerai ma réponse ce soir. Pour l'instant c'est 50-50.»* annonce-t-elle à l'agence locale. Il faut qu'elle se décide vite : l'excursion a lieu le lendemain. Mais Maurice tombe malade entre temps et quand il se réveille, surprise ! il se réveille dans un 4x4 et voit au loin une oasis et des touristes qui se promènent. Il réalise ceci : Jennifer a choisi l'une des deux excursions la veille et inconscient (ou oubli dû à la fièvre) on l'a fait monter dans un véhicule au petit matin. Il n'a aucune notion de l'heure qu'il est, matin, midi ou soir, ni de l'excursion choisie par Jennifer. On suppose que toutes les visites d'oasis ont la même durée D. *Quelle est la probabilité que Jennifer ait choisi la première formule du point de vue de Maurice ?*

La Belle au bois dormant : Les expérimentateurs endorment **Aurore** le dimanche soir, puis font un tirage au sort avec une pièce non truquée. Face : on réveille Aurore le lundi et les expérimentateurs ont un entretien avec elle. Pile : on réveille Aurore le lundi et les expérimentateurs ont un entretien avec elle, puis ils l'endorment à nouveau avec un traitement qui lui fait oublier la journée du lundi ; les expérimentateurs la réveillent une seconde fois le mardi et ont un entretien avec elle. *Du point de vue d'Aurore, à chaque entretien, quelles sont les probabilités de pile et face ?*

Dans ces quatre histoires, on constate que l'on part d'un vecteur de probabilités $\overrightarrow{P(A)} = (1/2 \quad 1/2)^t$. A l'aide de la notion d'univers, de graphes en treillis ou arborescent et de notation matricielle, on vérifiera si l'on observe un glissement de cette valeur.

2 Leurs Univers : la Rigueur

On résumera ici la notion *d'espace de probabilité* avec les trois éléments suivants : un ensemble d'événements (que l'on appellera *univers*), le tirage au sort et les probabilités associées à chacun de ces événements. Dans un premier temps, il vaut mieux considérer deux univers distincts, celui de départ et celui d'arrivée. Pour Aurore, l'univers de départ est celui de pile ou face. On note les événements {e1, e2}. Les trois réveils (lundi matin avec face, lundi matin avec pile et mardi matin avec pile) sont notés {f1, f2, f3} et forment un second univers. Il est important de voir que **l'on obtient les mêmes notations pour les trois autres histoires et donc les mêmes démonstrations**. Imaginons une simulation faite par ordinateur (c'est un excellent TD pour un étudiant). Il vérifiera qu'il a correctement programmé les séquences.

Solution avec un graphe en treillis (1) ; on relie les deux ensembles d'événements de cette manière :

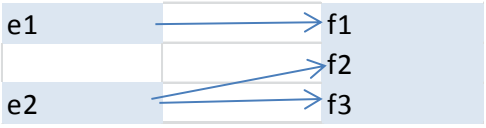

Figure 1 : Graphe en treillis entre les deux univers.

Puis en imaginant le processus effectué N fois, le tableau ci-dessous montre les différentes fréquences des événements croisés $e_x \cap f_y$:

	e1	**e2**
f1	N/2	0
f2	0	N/2
f3	0	N/2

Figure 2 : Comptage des événements croisés pour le graphe en treillis (1).

En normalisant, c'est-à-dire diviser ici par N/2, on obtient les différentes probabilités croisées $\mathrm{P}(\mathrm{e}_x \cap \mathrm{f}_y) = \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \\ 0 & 1/3 \end{pmatrix}$

En faisant la somme sur chaque ligne, on obtient directement les probabilités sur les événements f_y sous forme de vecteur $\overrightarrow{P(\mathrm{f})} = (1/3 \quad 1/3 \quad 1/3)^t$. En faisant la somme sur les colonnes, on obtient $\overrightarrow{P(\mathrm{e})} = (1/3 \quad 2/3)^t$.

Pour les probabilités conditionnelles, on applique $P(e_x|f_y) = P(e_x \cap f_y)/P(f_y)$. On obtient la matrice de probabilités conditionnelles $P(e_x|f_y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$. De façon similaire, on obtient $P(f_y|fe_{yx}) = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \\ 0 & 1/2 \end{pmatrix}$.

Solution avec un graphe en treillis (2) ; examinons ce qui se passe si l'on change un tant soit peu l'univers considéré. Par exemple, en regroupant f1 ∪ f2 = f'1, on obtient les fréquences suivantes :

	e1	**e2**
f'1 = f1 ∪ f2	N/2	N/2
f3	0	N/2

Figure 3 : Comptage des événements croisés pour le graphe en treillis (2).

En normalisant on obtient les différentes probabilités croisées $\mathrm{P}(\mathrm{e}_x \cap \mathrm{f}_y) = \begin{pmatrix} 1/3 & 1/3 \\ 0 & 1/3 \end{pmatrix}$. On obtient $\overrightarrow{P(\mathrm{f})} = (2/3 \quad 1/3)^t$, qui change de façon normale. Les nouvelles probabilités conditionnelles deviennent $P(e_x|f_y) = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix}$ et $P(f_y|e_x) = \begin{pmatrix} 1 & 1/2 \\ 0 & 1/2 \end{pmatrix}$. Mais on vérifie que $\overrightarrow{P(\mathrm{e})} = (1/3 \quad 2/3)^t$ ne change pas de valeurs.

Solution avec un graphe en arbre : une troisième méthode est possible en utilisant un schéma arborescent plus classique. On doit adjoindre le mardi matin pour face, mais le développement montrera que cela ne constitue pas une gêne. Pour les trois autres problèmes on obtient la même situation :

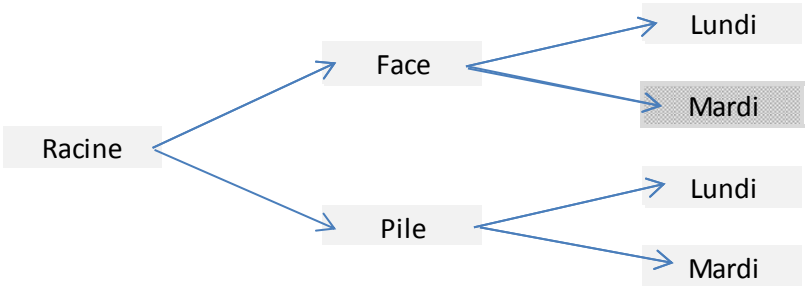

Figure 4 : Graphe arborescent pour l'ensemble des événements.

Avec N simulations on obtient 2N événements :

	e1	**e2**
f'1 = f1 ∪ f2	N/2	N/2
f3	N/2	N/2

Figure 5 : Comptage des événements croisés pour le graphe arborescent.

Cette méthode aboutit à la matrice de probabilités $\mathrm{P}(\mathrm{e}_x \cap \mathrm{f}_y) = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$.

Comme pour les deux démonstrations précédentes on obtient de nouvelles probabilités $\overrightarrow{P(\mathrm{f})} = (1/2 \quad 1/2)^t$. En faisant la somme sur les colonnes, on obtient $\overrightarrow{P(\mathrm{e})} = (1/2 \quad 1/2)^t$. Les nouvelles probabilités conditionnelles deviennent $P(e_x|f_y) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ et $P(f_y|e_x) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$.

Mais si l'on ne prend en compte que les événements en gris sur les tableaux ci-dessus, c'est-à-dire que les réveils lors de l'expérimentation, on obtient un événement r qu'Aurore (Marcel, Fabienne ou Maurice) doivent prendre en compte. Le calcul des nouvelles probabilités conditionnelles donnent $\overrightarrow{P(\mathrm{f}|\mathrm{r})} = (2/3 \quad 1/3)^t$ et $\overrightarrow{P(e|\mathrm{r})} = (1/3 \quad 2/3)^t$. Cette dernière relation est une autre écriture des $\overrightarrow{P(\mathrm{e})}$ vus précédemment. On peut vérifier que l'on peut créer autant d'univers que l'on veut en ajoutant des lignes et des colonnes en gris hachuré sur les tableaux vus plus haut, on ne change pas $\overrightarrow{P(e|\mathrm{r})} = (1/3 \quad 2/3)^t$. Mais surtout ici on observe un « glissement » de P(A). L'expérimentateur observe (1/2 ; 1/2) et Aurore (1/3 ; 2/3). Appliquer la formule de Bayes donnerait le même résultat.

3 L'Analyse des Résultats

Ces problèmes reposent sur des mécanismes similaires : les probabilités de départ passent de {1/2 ; 1/2} à {1/3 ; 2/3}. Mais on constate aussi que pour le Biathlon on part d'une observation des tirs sur les cibles. Pour les oasis et le morceau de jazz, un choix fait par respectivement la jeune touriste et l'amateur de musique. Pour la Belle, un tirage au sort. Mais la suite de chacune de ces petites histoires repose sur des événements qui sont de natures différentes : un passage sur une piste de pénalité, la présence près d'un oasis, l'arrivée de la jeune fille et le réveil d'Aurore. Effectivement, on a au début et à la fin de chaque histoire des ensembles d'événements bien distincts. Aucun lien entre un tir à la carabine et la présence du père au bon moment ! En fait dans tous ces cas, on a démontré qu'ici $\overrightarrow{P(A)}$ **change de valeur si on le change d'univers Ω**.

De façon similaire, on dira que dans chacun des problèmes, les observateurs ne sont pas dans des univers identiques et en conséquence ne voient pas le même vecteur $\overrightarrow{P(e)}$:

Enigme	« Demiste »	« Tiériste »
Biathlon	Entraîneur	Père du sportif
Take Five	Père	Fille
Oasis	Jeune fille	Jeune homme
Belle au bois dormant	Examinateurs	Aurore

En réalité dans la littérature on appelle « demistes » les auteurs qui ne considèrent pas valable le *glissement* de P(A). Ici on considère que l'on a quatre problèmes avec la présence dans chacun d'eux de deux univers distincts, le premier est « demiste » et le second est « tiériste », suivant les événements pris en considération par les observateurs :

Enigme	Événements « Demistes »	Événements « Tiéristes »
Biathlon	Résultat du tir	Le père voit son fils
Take Five	Choix aléatoire du père	La fille entend la musique

Oasis	Choix aléatoire de la jeune fille	Le jeune homme voit une oasis
Belle au bois dormant	Choix aléatoire des examinateurs	Aurore est réveillée pendant l'expérimentation

On note la présence d'événements dont la prise en compte ne modifierait pas le calcul, juste imposer l'emploi d'univers avec des événements supplémentaires et l'utilisation des probabilités conditionnelles :

Enigme	Événements écartés de l'énoncé
Biathlon	Le père arrive trop tard pour voir son fils
Take Five	La fille arrive après la fin de la version courte
Oasis	Le jeune homme se réveille entre deux oasis
Belle au bois dormant	Aurore se réveille le mardi matin hors du cadre de l'expérimentation

On continuerait à observer le *glissement* de $(1/2 ; 1/2)^t$ vers $(1/3 ; 2/3)^t$.

4 Variantes : des Univers à entropies H différentes

Modifions *Take Five*. Mathilde, la mère de Fabienne entend Pierre lancer l'air de jazz. Elle suppose que sa fille va bientôt arriver. Elle envisage les quatre possibilités vues plus haut comme équiprobables et les probabilités de voir soit la version courte soit la version longue sont $(1/2 ; 1/2)^t$. Seconde variante : Fabienne arrive avant la fin du morceau et demande à son père si c'est la version courte[i]. Celui-ci a oublié son choix et se retrouve dans la situation de Fabienne avec 3 cas possibles sur les quatre. Le vecteur devient $(1/3 ; 2/3)^t$. Et dernière variante : Fabienne arrive avant la fin du morceau et demande à son père si c'est la version courte. Celui-ci répond : *« Je ne sais pas, mais 5 minutes ne sont pas encore écoulées depuis le début du morceau. »*. Ceci élimine un cas et il n'en reste plus que deux. Le vecteur redevient $(1/2 ; 1/2)^t$!

On peut proposer un critère de **choix de l'univers** Ω à prendre en compte : celui ayant **la plus faible entropie** $\mathbf{H} = \sum_i p_i \log_2(p_i)$. Cette méthode a fait ses preuves en théorie de l'information. Dans les trois variantes ci-dessus, les trois entropies décroissent ; on obtient 2, 1,6 puis 1 *Shannons*. Si cela est possible, on choisira donc le denier scénario. (Néanmoins on note que même si l'entropie H diminue, le vecteur de probabilités ne converge nullement de façon monotone vers une valeur évidente. Mais ce point est un autre débat.)

5 Généralisation

Les paradoxes gagnent en clarté si l'on trouve une analogie basée sur un tirage avec des urnes. Soit deux urnes A et B ayant respectivement 1 et 2 boules blanches et un nombre *indifférent* de boules noires. On mélange le contenu des urnes, puis on effectue un tirage au sort. *Si l'on sort une boule blanche, quelle est la probabilité qu'elle vienne de A ou B ?*

On résume le problème avec le vecteur de probabilités du choix des urnes : $\vec{P} = (1/2 ; 1/2)^t$ et un vecteur de scalaire dénombrant le nombre de boules blanches des urnes : $\vec{N} = (1 ; 2)^t$. On peut ajouter autant de boules noires, ceci ne change pas le résultat. Le vecteur $\overrightarrow{P'}$ vu par Aurore, Marcel, Fabienne et Maurice est (le numérateur note la multiplication terme à terme et le dénominateur le produit scalaire) :

$$\overrightarrow{P'} = \frac{\vec{P} *. \vec{N}}{\vec{P} . \vec{N}}$$

Pour les cas simples du *Biathlon*, de *Take Five*, des *Oasis* et de la *Belle au bois dormant*, on obtient :

$$\overrightarrow{P'} = \frac{\binom{1/2}{1/2} *. \binom{1}{2}}{\binom{1/2}{1/2} . \binom{1}{2}} = \frac{\binom{1/2}{1}}{3/2} = \binom{1/3}{2/3}$$

Si l'on remplace respectivement $\vec{P}$ et $\vec{N}$ par des vecteurs plus complexes, la formule qui permet de calculer $\overrightarrow{P'}$ ne change pas de forme. Changer les pièces de monnaie par des dés, augmenter les pénalités du biathlon ou le nombre d'excursions vers les oasis... et appliquer la même formule. La formule rigoureuse est simple à démontrer. Par exemple, 6 urnes de respectivement 2, 4, 6, 7, 9 et 10 boules dont le choix est fait par un premier examinateur suivant le vecteur : $\vec{P} = (0{,}1\ 0{,}1\ 0{,}1\ 0{,}2\ 0{,}2\ 0{,}3)^t$. En appliquant la formule, le second examinateur obtient : $\vec{P} = 1/74\ (2\ 4\ 6\ 14\ 18\ 30)^t$.

6 Ce qu'il faut retenir

Le problème de la *Belle au bois dormant* peut être transposé à quantité d'autres tout aussi amusants. La méthode proposée ici est appliquée à quantité de domaines, dont la Théorie de l'information. Plusieurs démonstrations sont possibles. Néanmoins, on aboutit toujours à ce constat : un premier observateur voit un vecteur de probabilités et un second, disons plus en aval, en voit un autre. **Deux observateurs ne voient le même vecteur de probabilités $\overrightarrow{P(A)}$ que s'ils sont dans le même Espace de probabilités.** Si un choix est possible **on choisira celui présentant la plus faible Entropie H.** Au-delà des aspects de vulgarisation ludique, on observe que la simulation du problème est accessible à beaucoup si certaines bases sont assimilées de façon saine.

7 Exemple de références bibliographiques

Jean-Paul Delahaye [1] [2] aborde le problème à l'aide notion de filtres et de loupes pour illustrer les différents emplois de la formule de Bayes.

Laurent Delabre [3] propose une compilation des réponses apportées au problème de la *Belle au bois dormant* et souligne la controverse qui reste sur ce sujet.

Alexandru Spătaru [4] propose un cours détaillé dont on a repris et transposé les canaux d'information bruités.

Bibliographie

[1] Jean-Paul Delahaye (2008), Au pays des Paradoxes, Belin – Pour la Science, Chapitre 28.

[2] Jean-Paul Delahaye (2003), La Belle au bois dormant, la fin du monde et les extraterrestres, Pour la science. http://www.anthropic-principle.com/preprints/delahaye2.pdf.

[3] Laurent Delabre (2009), La Belle au bois dormant : débat autour d'un paradoxe.

[4] Alexandru Spătaru (1987), Fondement de la théorie de la transmission de l'information, Presses Polytechniques Romandes.

[i] Les énigmes sont de ma part, mais cette variante et la suivante auxquelles j'ai adapté mes démonstrations m'ont été suggérées par L. Delabre que je remercie.