Tentative d'identification de paradoxes latents dans l'application des probabilités conditionnelles

Léo Gerville-Réache, Vincent Couallier

TENTATIVE D'IDENTIFICATION DE PARADOXES LATENTS DANS L'APPLICATION DES PROBABILITES CONDITIONNELLES

Léo Gerville-Réache[1] & Vincent Couallier[2]

[1]*Université de Bordeaux, CNRS, UMR 5251, France, leo.gerville@u-bordeaux2.fr*
[2]*Université de Bordeaux, CNRS, UMR 5251, France, vincent.couallier@u-bordeaux2.fr*

Résumé. Depuis que la logique existe, les paradoxes coexistent. Pour beaucoup d'entre eux, l'énoncé est suffisamment simple pour que le « commun des mortels » comme le spécialiste le plus aguerri puisse proposer sa solution. Dans cette communication, nous revenons sur quelques paradoxes probabilistes à l'énoncé très accessible et proposons un regard statistique sur les causes de divergence des solutions proposées. Il semble que la polysémie du concept de probabilité conditionnelle soit souvent au cœur des débats.

Mots-clés. Paradoxe probabiliste, probabilité conditionnelle, estimateur, intervalle de confiance, les Sophies, les deux cravates, la Belle au bois dormant, l'apocalypse.

Abstract. Since the logic is, paradoxes are. For many of them, the statement is simple enough for the "common man" as for the most seasoned professional can provide his solution. In this paper, we review some very accessible probabilistic paradoxes and offer an outside point of view of the causes of multiple solutions. It seems that the polysemy of the concept of conditional probability is in the heart of the problem.

Keywords. Probabilistic paradox, conditional probability, estimator, confidence interval, the Sophies, the two ties, the Sleeping Beauty, Doomsday argument.

1 Introduction

L'analyse des « paradoxes » fait partie des problèmes les plus passionnants des sciences. De l'extérieur, on assiste à des discussions vigoureuses entre scientifiques, de l'intérieur, lorsqu'on se prend au jeu, notre esprit est bouleversé entre nos certitudes démontrées et celles, également démontrées, de nos contradicteurs. On se demande souvent comment il est possible que le temps d'attente d'un consensus soit parfois de plusieurs dizaines d'années. A posteriori, lorsqu'un paradoxe est résolu, ou fait clairement consensus, on est souvent étonné, voire stupéfait, de la simplicité de la solution.

Les nombreux paradoxes probabilistes sont source d'une production scientifique et non scientifique abondante. Parmi les plus célèbres, on peut citer le paradoxe de St-Petersbourg, le paradoxe du Monty Hall ou encore celui de l'Apocalypse. Les plus déstabilisants, que vous soyez spécialiste ou non, sont sûrement ceux où le calcul d'une probabilité conditionnelle est sous-jacente.

Les Sophies (Source J.P. Delahaye) : *Problème 1 : Une famille de deux enfants a au moins une fille. Quelle est la probabilité P_1, pour que cette famille ait deux filles ? Problème 2 : Une famille de deux enfants a au moins une fille qui s'appelle Sophie. Quelle est la probabilité P_2 pour que cette famille ait deux filles ?* Réponses proposées : on suppose que la probabilité qu'un enfant soit une fille vaut 1/2 et on néglige la question des jumeaux. Pour le problème 1 : les uns disent que $P_1=1/3$, les autres $P_1=1/2$. Nous disons que les deux sont possibles. Pour le problème 2 : les uns disent que la seule solution est $P_2 \approx 1/2$. Nous disons que 1/3 est également une réponse possible.

La Belle au bois dormant (Source P. Franceschi) : *Des chercheurs se préparent à vous endormir. Pendant les deux jours que durera votre sommeil, ils vous réveilleront brièvement soit une soit deux fois, selon le résultat du lancer d'une pièce de monnaie équilibrée (face: une fois; pile: deux fois). Après chaque réveil, ils vous endormiront à nouveau à l'aide d'une drogue qui vous fera oublier ce* *dernier réveil. Lorsque vous êtes réveillé, à quel degré devez-vous croire que le résultat du lancer de la monnaie est face?* Réponses proposées : les uns disent que la Belle doit répondre 1/2, les autres disent qu'elle doit répondre 1/3. Nous disons que les deux sont possibles.

Les deux cravates (Source J.P. Delahaye) : *Monsieur A propose à monsieur B que celui qui a la plus longue cravate la donne à l'autre. Monsieur A et monsieur B ont-ils intérêt à jouer ?* Réponses proposées : les uns disent qu'il faut jouer car l'espérance de gain est positive pour les deux joueurs, d'autres disent que le problème n'a pas de solution probabiliste (principe d'indifférence). Nous disons que le jeu est bien à somme nulle.

L'apocalypse (Source Wikipédia) : *Deux hypothèses sont en concurrence : la théorie A affirme que l'humanité disparaitra en 2150, et la théorie B affirme que cela sera beaucoup plus tard. Toujours selon l'hypothèse A, un humain sur 10 aura connu l'an 2000, et l'humanité aura compté 50 milliards d'individus depuis son origine. L'hypothèse B affirme qu'un humain sur 1 000 aura connu l'an 2000, et alors l'humanité aura compté le chiffre astronomique de 5 000 milliards de personnes. La théorie A semble la moins probable : on lui associe la probabilité a priori de 1 %, tandis que la théorie B bénéficie d'une probabilité de 99 %. Maintenant, considérons un événement E, par exemple « un individu fait partie des 5 milliards d'individus qui ont connu l'an 2000 ». On peut se demander « Quelle est l'hypothèse la plus probable, sachant cet événement ? ».* Réponses proposées : les uns disent que la probabilité de la théorie *A* est passée à 50,25%, les autres disent qu'elle est toujours de 1%. Nous disons que les deux réponses sont possibles.

L'intervalle de confiance (**Source B. Lecoutre**) *Quelle est la probabilité qu'un paramètre inconnu appartienne à la réalisation d'un intervalle de confiance à 95%.* Réponses proposées : les uns disent que, dans un cadre bayésien, la probabilité pourra être de 95%, les autres disent que, dans un cadre fréquentiste, cette probabilité vaut 0 ou 1. Nous disons que, dans un cadre fréquentiste, cette probabilité vaut également 95%.

Que vous connaissiez ou pas ces paradoxes, il est vraisemblable que vous ayez un avis, une conviction voire une certitude sur chacun d'eux. Il convient cependant, à l'abord d'un paradoxe, de conserver une extrême prudence. Les vérités se transforment parfois en conviction puis en doute pour finir en prise de conscience.

2 Qu'est-ce qu'une probabilité conditionnelle ?

D'un point de vue mathématique, une probabilité conditionnelle est une expression des plus simples. En théorie des probabilités, la probabilité conditionnelle d'un événement *A*, sachant qu'un autre événement *B* de probabilité non nulle s'est réalisé (ou probabilité de *A*, sachant *B*) est le nombre noté $\mathbf{P}(A|B)$ défini par $\mathbf{P}(A|B) = \mathbf{P}(A \cap B)/\mathbf{P}(B)$. On a $\mathbf{P}(A|B) = \mathbf{P}(B|A)\mathbf{P}(A)/\mathbf{P}(B)$ et $\mathbf{P}(A) = \mathbf{P}(A|B)\mathbf{P}(B) + \mathbf{P}(A|\bar{B})\mathbf{P}(\bar{B})$. Si *A* et *B* sont indépendants, on a : $\mathbf{P}(A|B) = \mathbf{P}(A)$. La simplicité de ces écritures cache deux risques majeurs dans leurs applications. Le premier, bien connu est lié à l'ambigüité sur le schéma probabiliste sous-jacent à la réalisation des événements *A* et *B* (le paradoxe de Bertrand par exemple). Une autre source d'ambigüité, visiblement méconnue, est liée à l'interprétation même de **P**(*A*|*B*). Afin de bien visualiser la problématique liée à l'interprétation de **P**(*A*|*B*), considérons les variables aléatoires suivantes :

Soit $\mathscr{X}$ la variable aléatoire telle que : $\begin{cases} \mathscr{X} = \mathbf{P}(A|B) \text{ avec probabilité } \mathbf{P}(B) \\ \mathscr{X} = \mathbf{P}(A|\bar{B}) \text{ avec probabilité 1-}\mathbf{P}(B) \end{cases}$

Soit $\mathscr{B}$ la variable aléatoire telle que : $\begin{cases} \mathscr{B} = B \text{ avec probabilité } \mathbf{P}(B) \\ \mathscr{B} = \bar{B} \text{ avec probabilité 1-}\mathbf{P}(B) \end{cases}$

Alors $\mathbf{P}(A|\mathscr{B})$ est une variable aléatoire dont l'espérance vaut :

$$\mathbf{E}\left(\mathbf{P}\left(A|\mathscr{B}\right)\right)=\mathbf{E}(\mathscr{X})=\mathbf{P}\left(A|B\right)\mathbf{P}(B)+\mathbf{P}\left(A|\bar{B}\right)\mathbf{P}(\bar{B})=\mathbf{P}(A)$$

D'un point de vue statistique, il s'en suit que $\mathbf{P}\left(A|\mathscr{B}\right)$ est un estimateur sans biais de $\mathbf{P}(A)$ et que $\mathbf{P}\left(A|B\right)$ est une estimation de $\mathbf{P}(A)$. On a également :

$$\mathbf{V}(\mathscr{X})=\mathbf{P}\left(A|B\right)^2\mathbf{P}(B)+\mathbf{P}\left(A|\bar{B}\right)^2\mathbf{P}(\bar{B})-\mathbf{P}(A)^2$$

NB : Si $\mathbf{V}(\mathscr{X})\neq 0$ alors $\mathbf{P}\left(A|B\right)$ et $\mathbf{P}\left(A|\bar{B}\right)$ sont des réalisations différentes et possibles de $\mathscr{X}$ et donc des estimations différentes et possibles de $\mathbf{P}(A)$. Ainsi, sachant *B*, $\mathbf{P}(A)$ ne « glisse » pas vers $\mathbf{P}\left(A|B\right)$ mais $\mathbf{P}\left(A|B\right)$ est simplement une estimation de $\mathbf{P}(A)$. Si, en théorie des probabilités, l'indépendance de *A* et *B* se caractérise par $\mathbf{P}\left(A|B\right)=\mathbf{P}(A)$, l'exemple suivant montre que le point de vue statistique peut s'avérer plus complexe.

Soit $\mathscr{U}$ et $\mathscr{V}$ deux variables aléatoires réelles indépendantes de même loi uniforme sur l'intervalle $[a,b]$. On a :

$$\mathbf{P}(\mathscr{V}<\mathscr{U})=1/2 \text{ et } \mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)=\mathbf{P}(\mathscr{V}<u)=(u-a)/(b-a).$$

Comme nous l'avons vu, $(u-a)/(b-a)$ est donc l'estimation de $\mathbf{P}(\mathscr{V}<\mathscr{U})$ sachant la réalisation de l'événement « $\mathscr{U}=u$ ». Si $u\neq(a+b)/2$, on a $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)\neq\mathbf{P}(\mathscr{V}<\mathscr{U})$ ce qui signifie clairement que les événements $\mathscr{V}<\mathscr{U}$ et $\mathscr{U}=u$ ne sont pas indépendants.

Mais que peut-on dire de $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)$ lorsque les réels *a* et *b* sont inconnus ?

Il est clair qu'on a toujours $\mathbf{P}(\mathscr{V}<\mathscr{U})=1/2$. On a également $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}\right)=(\mathscr{U}-a)/(b-a)$ et donc $\mathbf{E}\left(\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}\right)\right)=\mathbf{E}\left((\mathscr{U}-a)/(b-a)\right)=1/2$. Mais en l'absence de la connaissance des valeurs de *a* et *b*, $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)$ est inconnue. Il s'en suit que $\mathbf{E}\left((\mathscr{U}-a)/(b-a)\right)=1/2$ est la seule estimation (issue d'un estimateur sans biais) pour $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)$.

En conséquence, quel que soit *u*, $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)=1/2=\mathbf{P}(\mathscr{V}<\mathscr{U})$.

Nous pouvons donc conclure que la connaissance de la réalisation $\mathscr{U}=u$ n'est « utilisable » que si les réels *a* et *b* sont connus. Bien que les événements $\mathscr{V}<\mathscr{U}$ et $\mathscr{U}=u$ ne soient pas indépendants, si les entiers *a* et *b* sont inconnus $\mathbf{P}\left(\mathscr{V}<\mathscr{U}|\mathscr{U}=u\right)$ sera néanmoins estimée par $\mathbf{P}(\mathscr{V}<\mathscr{U})$.

Ces quelques éléments d'interprétation statistique des probabilités conditionnelles permettent de proposer un regard différent sur les cinq « paradoxes » présentés en introduction.

3 Analyse succincte des cinq « paradoxes »

Pour les Sophies, dans le problème 1, la prémisse est « Une famille de deux enfants a au moins une fille » et la question est « Quelle est la probabilité P_1, pour que cette famille ait deux filles ? ». Pour mettre cet énoncé sous forme d'une probabilité conditionnelle de la forme **P**(*A*|*B*), il nous faut d'abord définir les évènements *A* et *B*. L'événement *A* semble clair : *A* = « cette famille a deux filles ». Mais que prend-on pour *B* ? Le choix de *B* dépend directement de la façon dont l'information « une famille de deux enfants a au moins une fille » a été obtenue.

Ici, au moins deux schémas semblent possibles :

1) soit on a sélectionné « au hasard » une famille dans l'ensemble des familles de deux enfants et il s'est avéré que cette famille a au moins une fille (on s'intéresse à cette famille).

2) soit on a sélectionné « au hasard » un enfant dans l'ensemble des enfants appartenant aux familles de deux enfants et il s'est avéré que cet enfant était une fille (on s'intéresse à la famille de cette fille).

Ces deux schémas aboutissent bien à dire « une famille de deux enfants a au moins une fille ». La prémisse ne définit donc pas ici de façon unique le schéma probabiliste sous-jacent.

NB : dans les deux cas, **P**(FF|F)=**P**(F|FF)**P**(FF)/**P**(F) avec **P**(FF)=1/4 et **P**(F|FF)=1 mais dans le cas (1), **P**(F)=**P**(FG)+**P**(GF)+**P**(FF)=3/4 et dans le cas (2), **P**(F)=1/2**P**(FG)+1/2**P**(GF)+**P**(FF)=1/2. Ainsi dans le cas (1), on a **P**(FF|F)=1/3 alors que dans le cas (2), on a **P**(FF|F)=1/2. L'origine de la situation « paradoxale » vient du fait qu'un tiers des familles de deux enfants, ayant au moins une fille a deux filles mais que dans les familles de deux enfants, la moitié des filles a une sœur.

Le problème 2 (des Sophies), diffère du problème 1 par l'ajout des mots « qui s'appelle Sophie ». Ici également, au moins deux schémas probabilistes sont possibles :

1) soit on a sélectionné « au hasard » une famille dans l'ensemble des familles de deux enfants ayant au moins une Sophie (on s'intéresse à cette famille).

2) soit on a sélectionné « au hasard » une famille dans l'ensemble des familles de deux enfants ayant au moins une fille et il s'avère que cette fille se prénomme Sophie (on s'intéresse à cette famille).

La probabilité qu'une fille s'appelle Sophie n'est pas spécifiée. Si l'on suppose que le choix d'un prénom féminin est le résultat d'un tirage aléatoire équiprobable sur le grand ensemble des prénoms féminins existants, alors dans le cas (1), on trouve $\mathbf{P}(A|B) \approx 1/2$ car la probabilité du prénom Sophie est la seule à être « doublée » chez les familles avec 2 filles. Dans le cas (2), on trouve $\mathbf{P}(A|B) = 1/3$ car la probabilité du prénom obtenu est « doublée » au même titre que les probabilités de tous les prénoms féminins que l'on aurait pu obtenir ainsi chez les familles ayant 2 filles.

NB : on note π_s= **P**(S|F), la probabilité qu'une fille s'appelle Sophie. On suppose qu'une famille n'a pas deux Sophies. Dans le schéma (1) on a **P**(FF)=1/3 et **P**(S|FF)=$2\pi_s(1-\pi_s)$ et **P**(FF|S)=$(1-\pi_s)/(2-\pi_s)$ car **P**(S)=**P**(S|FG)**P**(FG)+**P**(S|GF)**P**(GF)+**P**(S|FF)**P**(FF)=$\pi_s(2-\pi_s)/2$. Dans le schéma (2) on montre que l'égalité des probabilités des prénoms féminins conduit, via la formule des probabilités totales appliquée à **P**(FF) à **P**(FF)=**P**(FF|S)=1/3. On remarque également que si les probabilités sur les prénoms féminins ne sont pas égales, le fait qu'elles soient inconnues conduit également, dans le schéma (2) à l'estimation 1/3 pour **P**(FF|S).

Selon le choix des schémas probabilistes pour le problème 1 puis pour le problème 2, on conclura que l'information « qui s'appelle Sophie » apporte ou pas une information « utilisable ».

Pour l'apocalypse, le conditionnement est le fait que vous êtes l'observateur et que vous faites partie des 5 milliards d'individus qui ont connu l'an 2000. Posons le schéma expérimental suivant : on sélectionne, au hasard (avec les probabilités **P**(*A*) et **P**(*B*)) une théorie, puis on sélectionne au hasard équiprobable un humain dans l'ensemble des humains faisant partie de la théorie sélectionnée. Il s'avère que cet humain a connu l'an 2000 (événement *E*). Un calcul de probabilité conditionnelle conduit à estimer la probabilité de la théorie *A* par **P**(*A*|*E*)=50,25%.

Pour autant, **P**(*A*) la probabilité de la théorie *A* a-t-elle « changé », « glissé »… vers **P**(*A*|*E*) ? Non. Si l'on considère l'ensemble des individus concernés par l'énoncé, un très grand nombre d'entre eux (par exemple, ceux qui naitraient après 2150) donneraient une estimation pour la probabilité de la théorie *A*, **P**(*A*|*E'*)=0% qui sera statistiquement tout aussi juste.

Donc, pour certains individus, l'estimation sera de 50,25% et pour d'autres, elle sera par exemple de 0%. Il s'en suit qu'on peut avoir simultanément (a) et (b) ou encore simultanément (a) et (c) :

(a) La probabilité que l'humanité disparaisse en 2150 est **P**(*A*) = 1%.

(b) Un humain sélectionné par hasard, ayant connu l'an 2000 estimera $\mathbf{P}(A)$ par $\mathbf{P}(A|E)$ = 50,25%.

(c) Un humain, sélectionné par hasard né après 2150 estimera $\mathbf{P}(A)$ par $\mathbf{P}(A|E')$=0%.

Ce n'est pas le fait de connaître l'an 2000 qui permet de dire $\mathbf{P}(A|E)$ = 50,25%, c'est le fait d'avoir été sélectionné par hasard selon le schéma expérimental proposé.

Pour la Belle au bois dormant, la source de confusion réside sûrement dans l'ambigüité de la question posée à la Belle : « *Lorsque vous êtes réveillée, à quel degré devez-vous croire que le résultat du lancer de la monnaie est face* ? ». Dans quelle mesure la question évoque ici la « proportion de réveils » ou la « proportion d'expériences » ? Pour s'en convaincre, réalisons une expérience où on note X la variable aléatoire du jet de pièce Pile/Face, et définissons deux stratégies pour la Belle : répondre systématiquement « je pense que la pièce est tombée sur Pile », avec Y le nombre de fois où la Belle a raison ; ou répondre systématiquement « je pense que la pièce est tombée sur Face», avec Z le nombre de fois où la Belle a raison. Dans ce cas, une expérience peut fournir les deux résultats suivants :

Résultat du Pile/Face	Valeurs de Y	Valeurs de Z	Probabilité	Valeurs Y+Z	Valeurs Y/(Y+Z)	Valeurs Z/(Y+Z)
X=Pile	2	0	0.5	2	1	0
X=Face	0	1	0.5	1	0	1

On peut donc interpréter à *quel degré devez-vous croire que le résultat du lancer de la monnaie est face* ? par : (1) « quelle croyance avoir dans $Z > Y$ », c'est à dire que la Belle a plus de croyance que le lancer de la monnaie ait donné face ? Or, ici $\mathbf{P}(Z>Y)=\mathbf{P}(Z<Y)=1/2$; ou (2) comparer le nombre moyen de réussites dans l'expérience (unique) quand on parie sur Pile et quand on parie sur Face. Or ici, $\mathbf{E}(Y)=1$ et $\mathbf{E}(Z)=1/2$. Comme $\mathbf{E}(Z)>\mathbf{E}(Y)$, on a moins à gagner en pariant sur Face.

En fait, le nombre de questionnements à la Belle à chaque expérience est $N=Y+Z$ et prend les valeurs 1 ou 2 avec probabilités 1/2. Ainsi, *à chaque expérience*, $Z/(Y+Z)$ donne la fréquence aléatoire du nombre de Face « vus » par la Belle. En moyenne (par expérience), $Y/(Z+Y)$ et $Z/(Y+Z)$ valent 1/2 (i.e. $\mathbf{E}(Y/(Z+Y))=\mathbf{E}(Z/(Y+Z))=1/2$). Par contre, $\mathbf{E}(Z)/\mathbf{E}(Z+Y)=1/3$! Ce dernier représente le gain moyen par réveil, au sens où, en moyenne, la Belle est réveillée 1,5 fois par expérience, et Z a pour moyenne 1/2 (par expérience). La probabilité de gagner à chaque pari (1/2) n'est pas égale à la proportion de paris gagnés (1/3). Ces techniques de changements de probabilité sont à la base de certaines méthodes de simulation de Monte Carlo (méthode du rejet, importance sampling).

Pour les deux cravates, l'argument conduisant à dire que les deux joueurs ont des espérances de gain positives est le suivant : Monsieur A a une chance sur deux de perdre sa cravate de longueur L_A et une chance sur deux de gagner une cravate de longueur L_B, plus grande que la sienne. *En moyenne*, le gain pour A *serait* donc de $0.5L_B-0.5L_A>0$. Le même raisonnement fait par Monsieur B conduirait ainsi à un jeu « gagnant-gagnant ». Posons les choses. Si Monsieur A décide de jouer :

- soit $L_A > L_B$ et dans ce cas, A perd sa cravate et B gagne L_A,
- soit $L_A < L_B$ et dans ce cas, A gagne L_B et B perd sa cravate.

Si Monsieur A décide de ne pas jouer, Monsieur A conserve sa cravate (et donc ne gagne rien). Sans connaissance initiale sur les longueurs des cravates, imaginons que L_A et L_B sont aléatoires et indépendantes, provenant de la même loi de tirage d'une cravate dans l'ensemble des cravates existantes pouvant prendre toute longueur (positive). Notons E, l'espérance de cette longueur. La première stratégie consistant à refuser le jeu donnera un gain nul à Monsieur A, $G_1=0$. La seconde stratégie, consistant à accepter le jeu donnera un gain négatif aléatoire $G_2=-L_A$, si $L_A > L_B$, avec la probabilité $\mathbf{P}(L_A > L_B)$ ou un gain positif aléatoire $G_2=L_B$ si $L_A< L_B$, avec la probabilité $P(L_A < L_B)$.

Si $\mathbf{P}(L_A=L_B)=0$, comme L_A et L_B sont de même loi de probabilité, on sait que $\mathbf{P}(L_A > L_B)= \mathbf{P}(L_A < L_B)=1/2$ et G_2 peut s'écrire :

$$G_2 = -L_A I_{\{L_A>L_B\}} + L_B I_{\{L_A<L_B\}}.$$

On peut donc choisir de jouer selon deux critères : $\mathbf{P}(G_2>G_1)=\mathbf{P}(G_2>0)>1/2$ ou $\mathbf{E}(G_2)>\mathbf{E}(G_1)=0$. Mais il est faux de mélanger espérance et aléatoire comme dans la phrase qui suit :

$$E(G_2) = -L_A \mathrm{P}\{L_A > L_B\} + L_B \mathrm{P}\{L_A < L_B\} = (L_B - L_A)/2 > 0 \text{ : ceci est faux !}$$

On a : $E\left(-L_A I_{\{L_A>L_B\}} + L_B I_{\{L_A<L_B\}}\right) = E\left(-L_A + (L_A+L_B) I_{\{L_A<L_B\}}\right) = -E + E\left((L_A+L_B) I_{\{L_A<L_B\}}\right) = -E + E = 0$

De même, on montre aisément que $\mathbf{P}(G_2>0)=1/2$. Le jeu est donc équilibré. Un argument est alors soulevé : «connaître la longueur de la cravate me fait-il changer d'avis ? ». Dois-je jouer si je sais que ma cravate fait $L_A=l_a$ cm ? La réponse dépend de façon surprenante du fait de connaitre en partie la loi de probabilité des longueurs. Si la loi est totalement inconnue, les probabilités conditionnelles (sachant $L_A=l_a$) ne changent pas. Si la loi est en partie connue, permettant de calculer $\mathbf{P}(L>l_a)$, on décidera de jouer si $\mathbf{P}(L>l_a)>1/2$. En effet, la variable G_2 prend les mêmes valeurs ($-L_A$ et L_B), mais avec les probabilités conditionnelles 1- $\mathbf{P}(L>l_a)$ et $\mathbf{P}(L>l_a)$. La probabilité d'un gain positif est alors $\mathbf{P}(\mathbf{G}_2>0)=\mathbf{P}(L_A<L_B|L_A=l_a)=\mathbf{P}(L_B>l_a)=\mathbf{P}(L>l_a)>1/2$.

Pour l'intervalle de confiance, si un échantillon permet de conclure qu'un intervalle de confiance à 95% pour un paramètre π est [0,58 ; 0,64] alors on ne pourrait pas écrire, sauf à mettre une loi de probabilité sur π, que : $\mathbf{P}\left(\pi \in [0,58; 0,64]\right) = 0,95$. Attention, la probabilité qui nous intéresse est : « La probabilité que π appartienne à la réalisation suivante d'un intervalle de confiance à 95% : [0,58 ; 0,64] ». Cette probabilité s'écrit : $\mathbf{P}\left(\pi \in \left[IC_{\inf}; IC_{\sup}\right] \middle| IC_{\inf} = 0,58 \cap IC_{\sup} = 0,64\right)$. La valeur de π étant inconnue, les événements $IC_{\inf} = 0,58 \cap IC_{\sup} = 0,64$ et $\pi \in \left[IC_{\inf}; IC_{\sup}\right]$ sont indépendants et il s'en suit que :

$$\mathbf{P}\left(\pi \in \left[IC_{\inf}; IC_{\sup}\right] \middle| IC_{\inf} = 0,58 \cap IC_{\sup} = 0,64\right) = \mathbf{P}\left(\pi \in \left[IC_{\inf}; IC_{\sup}\right]\right) = 0,95.$$

Ainsi, sans modélisation bayésienne de π, 95% est l'estimation statistique de $\mathbf{P}\left(\pi \in [0,58; 0,64]\right)$ compte tenu que $[0,58; 0,64]$ est la réalisation d'un intervalle de confiance à 95% pour π.

4 Conclusion

L'application et l'interprétation des probabilités conditionnelles est délicate. Les paradoxes présentés montrent des énoncés pouvant interpeler la raison jusqu'à l'insomnie. La démarche suivie dans ce travail a été double. Dans un premier temps, nous avons cherché des schémas probabilistes compatibles avec les énoncées permettant de formaliser, autant que possible, les réponses des uns et des autres. Cependant, il est apparu que la multiplicité des schémas probabilistes possibles n'expliquait pas, à elle seule, la compatibilité de certaines réponses proposées. Aussi, dans un deuxième temps, il est apparu qu'une interprétation statistique de l'application des probabilités conditionnelles devait être développée. Dans tous les cas, il est essentiel de ne pas oublier que les énoncés infimement polysémiques sont infiniment polémiques.

Bibliographie

[1] Delahaye, J.P. (2005). Le trésor et les Sophies. *Pour la science* N°336, 90-94.
[2] Delahaye, J.P. (2008). *Au pays des paradoxes*. Edition Belin - Pour la science.
[3] Emery, M. (2001). Quelques phénomènes curieux en probabilités et statistiques. *L'Ouvert*, N°104, 37-47.
[4] Franceschi, P. (2002). *Une application des n-univers à l'argument de l'apocalypse et au paradoxe de Goodman*. Thèse de l'université de Corse.
[5] Gerville-Réache, L., Couallier V., Bayle, F. (2011). Quatre approches pour l'estimation d'une probabilité de défaillance exponentielle à partir d'un unique essai zéro défaillance réussi, *Journée Fiabilité et Incertitude de la SFDS*, Paris, France.
[6] Lecoutre, B. (2005). Et si vous étiez un bayésien qui s'ignore? *Modulad* N° 32, 92-105.